

Unexploded Ordnance

A Critical Review of Risk Assessment Methods

Jacqueline MacDonald

Debra Knopman

J.R. Lockwood

Gary Cecchine

Henry Willis

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

RAND
Arroyo Center

Unexploded Ordnance

A Critical Review of Risk Assessment Methods

Jacqueline MacDonald

Debra Knopman

J.R. Lockwood

Gary Cecchine

Henry Willis

MR-1674-A

20041208 212 —

Prepared for the United States Army

**RAND
Arroyo Center**

Approved for public release; distribution unlimited

The research described in this report was sponsored by the United States Army under Contract No. DASW01-01-C-0003.

Library of Congress Cataloging-in-Publication Data

Unexploded ordnance : a critical review of risk assessment methods / Jacqueline MacDonald ... [et al.].
p. cm.
Includes bibliographical references.
“MR-1674.”
ISBN 0-8330-3432-4 (Paperback)
1. Explosive ordnance disposal. 2. Military bases—United States. 3. Military base closures—United States. 4. Hazardous substances—Risk assessment—United States. I. MacDonald, Jacqueline.

UF860.U55 2003
355.6'213'0973—dc21

2003012740

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

Cover design by Barbara Angell Caslon

© Copyright 2004 RAND Corporation

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from RAND.

Published 2004 by the RAND Corporation
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
201 North Craig Street, Suite 202, Pittsburgh, PA 15213-1516
RAND URL: <http://www.rand.org/>

To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

PREFACE

This report documents the findings of a study that examined methods for assessing the risks of unexploded ordnance (UXO) and munitions constituents on former military training land. The report focuses specifically on methods applicable to sites on closed, transferred, and transferring bases that are being or have already been converted to civilian uses. The Army has not yet arrived at a consensus with the Environmental Protection Agency (EPA), state agencies, and other concerned groups about what process should be used to assess risks and evaluate potential responses to sites contaminated with both UXO and munitions constituents. This report is the first to analyze in detail all of the approaches the Army has developed to assess risks of UXO sites. It also recommends how the Army can move forward with developing risk assessment protocols that will be acceptable to all of those involved at UXO sites.

The report should interest anyone involved in the transfer of military sites under the Base Realignment and Closure (BRAC) program and with management of the Formerly Utilized Defense Sites (FUDS). Although the report was written for the Army, it will also be of broad interest to the Department of Defense (DoD), the EPA, the Department of the Interior, state regulators, and citizen groups involved at BRAC and FUDS sites. In addition, aspects of this report should interest Army and DoD policymakers involved in the planning of possible future base closure rounds.

This work was sponsored by the Assistant Chief of Staff for Installation Management, United States Army. The research was conducted in the Military Logistics Program of the RAND Arroyo Center, a fed-

erally funded research and development center sponsored by the U.S. Army.

For more information on RAND Arroyo Center, contact the Director of Operations (telephone 310-393-0411, extension 6419; FAX 310-451-6952; e-mail Marcy_Agmon@rand.org), or visit the Arroyo Center's Web site at <http://www.rand.org/ard/>.

CONTENTS

Preface	iii
Figures	ix
Tables	xi
Summary	xiii
Acknowledgments	xxv
Abbreviations	xxvii
Chapter One	
INTRODUCTION	1
Study Tasks	2
Study Methods	3
How This Report Is Organized	4
Categories of UXO Sites	4
Number of UXO Sites	4
Sources and Types of UXO	5
Risks from UXO Sites	6
UXO Detection and Removal Process	9
Efforts to Standardize UXO Risk Assessment Processes	15
Chapter Two	
POTENTIAL FOR A RISK-BASED APPROACH	
TO UXO RESPONSE	19
What Is Risk Assessment?	20
History of Risk Assessment	21
Potential Applications of Risk Assessment to UXO Problems	24

Challenges of Using Risk Assessment for UXO Sites	24
Criteria for an Effective Risk Assessment Method	30
Conclusions	34
Chapter Three	
EVALUATION OF EXISTING METHODS	
FOR UXO RISK ASSESSMENT	35
Overview of Existing Methods	35
Intended Uses	36
Risk Sources, Data Requirements, and Output	37
Criteria for an Effective Method	37
Interim Range Rule Risk Methodology	38
Key Features	38
Description	39
Risk Calculation Method	39
Implementation Features	49
Communication Features	50
Conclusions	51
Ordnance and Explosives Cost-Effectiveness Risk Tool	51
Key Features	51
Description	52
Risk Calculation Method	53
Technical Models and Assumptions	56
Implementation Features	62
Communication Features	62
Conclusions	63
Risk Assessment Code	63
Key Features	63
Description	63
Risk Calculation Features	65
Implementation Features	66
Communication Features	66
Conclusions	66
Ordnance and Explosives Risk Impact Analysis	67
Key Features	67
Description	67
Risk Calculation Features	69
Implementation Features	69
Communication Features	71
Conclusions	71
Natural and Cultural Resources Bank	71

Key Features	71
Description	72
Risk Calculation Features	72
Implementation Features	76
Communication Features	77
Conclusions	78
Summary	78
Chapter Four	
OTHER FEDERAL RISK ASSESSMENT METHODS	81
Prioritizing Sites Based On Risk	84
Relative Risk Site Evaluation Primer	84
Hazard Ranking System	92
Environmental Restoration Priority System	95
Developing Quantitative Risk Estimates for Specific Sites	99
Risk Assessment Guidance for Superfund	100
Ecological Risk Assessment Guidance for Superfund	104
Probabilistic Risk Assessment (NRC, NASA, and FAA)	106
Safety Assessment for Explosion Risk	110
Assessing Multiple Risk End Points	112
Evaluating Uncertainty in Risk Estimates	116
Standardizing the Risk Assessment Process	118
Working with Stakeholders	119
Developing Trust Among Stakeholders	122
Conclusions	123
Chapter Five	
CONCLUSIONS AND RECOMMENDATIONS	125
Recommended Process for Prioritization	129
Process Description	129
Oversight by Technical Review Board and Advisory Committee	133
Two Lists, Based on Land Use	134
Recommended Site-Specific Risk Assessment Process	135
Description of Recommended Approach	136
Development of Event and Fault Trees	137
Oversight by a Technical Review Board	139
Summary of Recommendations	139
References	143

FIGURES

1.1.	60mm HE Mortar Found at Fort McClellan	6
1.2.	Searching for UXO at Fort Ord	10
1.3.	Example ROC Curve from Field Test of EMI Systems at Fort Ord (1.6-foot search radius)	12
1.4.	Example ROC Curve from Field Test of EMI Systems at Fort Ord (3.3-foot search radius)	13
1.5.	Example of a Process Proposed by State Regulators for Cleanup of a UXO Site	14
2.1.	Classic Example of Dose-Response Curves for Three Chemicals That Can Cause Cancer	28
2.2.	Plumes of Groundwater Contaminants at the Massachusetts Military Reservation	29
3.1.	Rules for Determining the Accessibility Factor in the IR3M Explosives Safety Risk Tool	41
3.2.	Rules for Determining the Overall Hazard Factor in the IR3M Explosives Safety Risk Tool	42
3.3.	Rules for Determining the Exposure Factor in the IR3M Explosives Safety Risk Tool	43
3.4.	IR3M Rules for Determining the Explosives Safety Risk of a UXO Site, on a Scale of A to E	44
3.5.	Input Factors and Rules for Determining the IR3M Value for Human Health Risk	46
3.6.	Input Factors and Associated Output for IR3M Ecological Risk Tool	49
4.1.	Flow Diagram Showing Inputs to and Output from the Relative Risk Site Evaluation	88
4.2.	RRSE Matrix for Assigning Site Scores	91
4.3.	RAGS Risk Assessment Steps	102

4.4.	Example Pathways for Exposure to Contaminants from a Waste Site, as Illustrated in RAGS	103
4.5.	Event Tree Showing Possible Scenarios After Three Spins of a Fair Coin	107
4.6.	Event Tree Showing Possible Scenarios After Flammable Liquid in Industrial Facility Catches Fire ..	108
4.7.	Example of a Fault Tree Used in Assessing Risks at the Tooele Chemical Agent Disposal Facility	114
4.8.	Contributors to the Average Public Fatality Risk Due To Chemical Weapons Storage and Disposal at the Tooele Chemical Agent Disposal Facility	115
5.1.	Recommended Two-Stage Prioritization Process	130
5.2.	Summary of Recommended Prioritization Process	132

TABLES

S.1. Summary of Needs for UXO Risk Assessment	xvi
S.2. Overview of UXO Risk Assessment Methods	xix
2.1. Questions Addressed by the Two Applications of Risk Assessment	25
2.2. Differences Between Explosion Risks and Munitions Constituents Risks	27
2.3. Criteria for Evaluating UXO Risk Assessment Methods	32
3.1. Features of UXO Risk Assessment Methods	36
3.2. Activities Resulting in Potential UXO Exposure	53
3.3. Environmental Data Used by OECert to Determine Homogeneity of UXO Sites or Sectors	54
3.4. Categories of Localized Sectors in OECert	55
3.5. Activities Considered by OECert when Estimating Exposure to UXO in Water Sectors	56
3.6. RAC Hazard Severity Values for Conventional Ordnance	64
3.7. RAC Matrix for Determining Priority for Action	65
3.8. Input Information for OERIA	68
3.9. Example of OERIA Output	70
3.10. NCRB Resource Values	73
3.11. NCRB Information Reliability Values	74
3.12. NCRB Cumulative Scores and Predictive Categories ..	75
4.1. Federal Risk Assessment Methods Reviewed	82
4.2. Calculation of CHF	89
4.3. Calculation of MPF for Groundwater	90
4.4. Weights Assigned to the Objectives of the DOE Environmental Management Program	97

4.5. Summary: Existing Federal Methods That Illustrate Solutions to Common Problems in Risk Assessment	124
5.1. Summary of UXO Risk Assessment Methods	126
5.2. Summary of Criteria for Evaluating UXO Risk Assessment Methods	131

SUMMARY

Unexploded ordnance (UXO) and munitions constituents¹ on former military bases in the United States are causing increasing concern. While civilian fatalities from UXO explosions on U.S. soil have been rare, the risk of such accidents could increase substantially as more closed bases are transferred from military to civilian control. Since the end of the Cold War, approximately 20 percent of major domestic military bases and many smaller ones have been closed and designated for eventual transfer to civilian ownership. Reflecting the growing concern about domestic UXO sites, the National Defense Authorization Act of 2002 directs the Department of Defense to inventory UXO sites, establish a new program element for UXO remediation, and assess progress to date on cleaning up UXO.

This report addresses one part of the process of cleaning up UXO and munitions constituents at domestic military installations: the assessment of risks associated with these contaminants. Risk assessment helps define the technical dimension of UXO problems. It provides a technical basis for setting priorities among sites and choosing among alternative cleanup strategies.

It is important to keep in mind that even the best-designed set of risk assessment methods will not resolve all the controversies that arise at UXO sites. Risk assessment can help to educate the participants in the decision process about the nature and magnitude of risk in-

¹The term “munitions constituents” refers to any materials originating from UXO or other munitions, including the chemical contaminants that result from their breakdown.

volved. However, the ultimate decision about how to respond to UXO must account for many other factors—including ethical concerns, socioeconomic issues, and costs—in addition to risk. The risk assessor's job is not to decide what risk is acceptable; it is to do the best possible job calculating the risk. Risk assessment can illuminate the nature of risks at UXO sites, but it cannot make people agree on what amount of risk is acceptable. Nonetheless, we believe that pursuing the recommendations below will lead to better-informed decisions about how to manage UXO sites.

In this report, we evaluate the adequacy of methods developed for UXO risk assessment, review the risk assessment methodologies of other federal agencies for possible application to UXO, and propose strategies for improving risk assessment methods for UXO sites.

MULTIPLE RISK ASSESSMENT METHODS ARE NEEDED

A single method for assessing risks at UXO sites will not suffice. Rather, the Army needs to develop different methods for different steps in the UXO risk assessment process and for different elements of UXO risk.

One set of risk assessment methods would establish priorities in the UXO response program. We call this type of method *programmatic prioritization risk assessment*. Such methods could inform decisionmakers about which installations and sites within installations pose the greatest risk and thus merit the most immediate attention. This type of information is useful for allocating financial and other resources, such as equipment and personnel.

The second set of risk assessment methods would provide detailed analyses of specific UXO-contaminated areas within installations. We call this type of method *site-specific risk assessment*. Site-specific risk assessment methods could provide quantitative information about the potential for harm to people living near UXO sites and to local ecosystems. They could also estimate the effectiveness of alternative UXO response options in reducing those risks.

Programmatic prioritization methods and site-specific methods would require different designs. Programmatic prioritization methods would serve as a coarse screen for large groups of sites; their

purpose would be to establish relative risk levels among sites. At the stage when prioritization is usually carried out, site data are often limited. In contrast, site-specific methods would serve as tools for understanding the details about how people and ecosystems might become exposed to UXO and the probable consequences of such exposures using information specific to the sites. Detailed data collection would be necessary for a comprehensive understanding of the risks. As a result of these differences, two risk assessment approaches are needed for UXO sites: one for programmatic prioritization and another for site-specific assessment.

Two sources of risk at UXO sites also must be considered: (1) risks from UXO explosions and (2) risks from munitions constituents that have leached into soil and water. These two hazards differ substantially in the nature of the threats they pose and in the reaction of stakeholders to them. For example, the consequence of a human accidentally detonating UXO is immediate and typically results in serious injury or death. In contrast, the consequence of a human exposure to munitions constituents is most likely chronic and increases the risk of illness only after prolonged exposure. As a result, the methods used to assess explosion risks—whether for establishing priorities or conducting detailed site investigations—will necessarily differ substantially from those used to assess munitions constituents risks.

Thus, one risk assessment method cannot meet all the Army's needs for UXO sites. Different methods are needed for site prioritization and for site-specific assessment. Within each of these methods, different approaches are required to evaluate munitions constituents and explosion risks. Table S.1 summarizes the needs for UXO risk assessment. The last column of the table identifies existing risk assessment methods that could be used or could serve as a model for developing a new method; the basis for this column is discussed below.

ADEQUACY OF AVAILABLE UXO RISK ASSESSMENT METHODS

The Army asked us to review five existing risk assessment methods that were designed specifically for UXO. We compared the attributes

Table S.1
Summary of Needs for UXO Risk Assessment

Use of Risk Assessment	Methods Required to Support Use	Example Questions Answered by Method	Applicable Existing Methods
Programmatic prioritization	Munitions constituents prioritization method	At which sites do munitions chemicals that have leached into soil and water pose the highest risks to public health?	EPA Hazard Ranking System Defense Environmental Restoration Program Relative Risk Site Evaluation Primer
UXO explosion prioritization method		At which sites do munitions chemicals in soil and water pose the highest risks to the environment? At which sites does the potential for accidental UXO detonation pose the highest risks to the public?	Risk Assessment Code (modified with stakeholder input)
Site-specific assessment	Munitions constituents site-specific assessment method	At which sites does the potential for accidental UXO detonation pose the highest risks to workers? What is the probability that those living near a specific UXO site will experience health problems (e.g., cancer, lead poisoning) due to exposure to munitions chemicals in local soil and water?	EPA Risk Assessment Guidance for Superfund EPA Ecological Risk Assessment Guidance for Superfund

Table S.1—continued

Use of Risk Assessment	Methods Required to Support Use	Example Questions Answered by Method	Applicable Existing Methods
	UXO explosion site-specific assessment method	<p>How and to what extent might munitions chemicals in soil and water damage the local ecosystem?</p> <p>How will the probability of adverse health consequences change due to specific remediation methods at sites?</p> <p>What is the probability that a person living near a given UXO site will be harmed by accidental UXO detonation?</p> <p>If all UXO items on the surface are cleared, what is the probability of a person being harmed by remaining, buried UXO?</p> <p>If detection devices are used to identify and clear UXO to a given depth, what is the probability of a person being harmed by any remaining UXO?</p> <p>If future land-use scenarios are changed, how will the probability of a person being injured by UXO change?</p>	Probabilistic risk assessment (used by the Nuclear Regulatory Commission, National Aeronautics and Space Administration, Army Chemical Stockpile Program, and others; details for UXO application would need to be developed)

of these methods to criteria necessary for a technically sound risk assessment. We developed these criteria based on a survey of risk assessment literature and consultations with leading national experts in risk assessment. We conclude that none of the five risk assessment methods fully answers the Army's needs, either for programmatic prioritization or for site-specific assessment. Table S.2 summarizes the results of our evaluation and the basis for this conclusion.

APPLICABLE CONCEPTS FROM RISK ASSESSMENT METHODS OF OTHER FEDERAL PROGRAMS

Our review of risk assessment approaches available from other federal programs indicates that some of them apply directly to assessing some (but not all) of the risks associated with UXO sites, as Table S.1 indicates. Many others are not directly applicable to UXO risk assessment, but they provide examples of approaches for addressing problems in risk assessment that the Army has encountered at UXO sites. Chapter Four describes a range of federal methods for analyzing uncertainty, involving stakeholders and gaining their trust, standardizing the risk analysis process, and considering multiple end points in risk assessment. We do not repeat the details here.

As Table S.1 shows, *existing methods from other programs can be applied directly to assessing the risks of munitions constituents in soil and water*. No new methods need to be developed for this purpose. The Defense Environmental Restoration Program Relative Risk Site Evaluation Primer and the Environmental Protection Agency's Hazard Ranking System, Risk Assessment Guidance for Superfund, and Ecological Risk Assessment Guidance for Superfund all meet needs for assessing risks from munitions constituents. The former two methods are well established for prioritizing sites according to risks posed by chemical contaminants that have dissolved in water, absorbed to soil, or dissolved in the air in spaces between soil grains. The latter two methods are well established for site-specific assessment of risks from contaminants in water or present in soil. Munitions constituents, when present in relatively dilute concentrations in soil and water, can be treated just as any other type of chemical contaminant in soil and water; they pose no unique risks, compared to other types of contaminants found at hazardous waste sites.

Table S.2
Overview of UXO Risk Assessment Methods

Method	Purpose	Pros	Cons	Summary Evaluation
Interim Range Rule Risk Methodology (IR3M)	Assess	Simple output	Output does not always correlate to risk Output can mask important risk information Decision rules not technically justified Basis for input values not justified Not always reproducible Does not address uncertainty Data requirements insufficient to reflect problem complexity Instructions unclear	Significant limitations; should not be developed further
Ordnance and Explosives Cost Effectiveness Risk Tool (OECDrt)	Developed to prioritize, but in practice used to assess	Comprehensive modeling of exposure process Analytical process used to determine explosion potential of different munitions Adaptable	Does not address munitions constituents risk Exposure models not validated Many exposure assumptions not justified Uncertainties not addressed Calculations not presented clearly Not easily communicated to stakeholders Lack of stakeholder involvement in developing exposure assumptions	Elements of the method (exposure models, UXO categorization method) might form part of future risk assessment method but would need much refinement

Table S.2—continued

Method	Purpose	Pros	Cons	Summary Evaluation
Risk Assessment Code (RAC)	Prioritize	Appears logically sound Assumptions clearly explained Reproducible Practical (data requirements suitable for purpose) Adaptable	Does not consider munitions constituents Does not address uncertainty Basis for some assumptions not provided Does not address munitions constituents risk Risk model relation to actual magnitude of site risk unknown Assumptions not explained Uncertainty not addressed Not reproducible Data requirements too minimal for use Results easily manipulated	Well suited for purpose, but only addresses explosion risk; assumptions may need to be modified with stakeholder input Has many limitations and should be discontinued
Ordnance and Explosives Risk Impact Analysis (OERIA)	Assess	Easy to use Adaptable	Focused exclusively on ecological risks Does not consider munitions constituents Assumptions not justified Uncertainties not addressed Instructions somewhat unclear	Meets need to identify UXO sites with regulatory requirements related to natural or cultural resources, but needs substantial further development and validation
Natural and Cultural Resources Bank (NCRB)	Prioritize	Appears to be reproducible Adaptable		

Table S.1 also indicates that existing methods from other programs could serve as models for developing new site-specific explosion risk assessment methods. The probabilistic risk assessment (PRA) approaches used by the Army Chemical Stockpile Disposal Program, the Federal Aviation Administration, the National Aeronautics and Space Administration, and the Nuclear Regulatory Commission provide representative examples. In particular, the fault tree and event tree analysis tools used in PRA seem most relevant. These tools are used for systematic mapping of the steps necessary to trigger acute events. The approaches are widely used not only in the federal government but also in industry for determining the probability of system failures and identifying the most important contributors to the risks of those failures. These tools could meet the need for a new method to assess site-specific UXO explosion risk.

No existing method is adequate for considering both explosion risk and munitions constituents risk when prioritizing sites. A new process will need to be developed for this purpose. Nonetheless, as Table S.1 shows, the Risk Assessment Code could be modified and combined with the Relative Risk Site Evaluation Primer or the Hazard Ranking System for the purpose of prioritization.

RECOMMENDATIONS

In summary, two separate processes are needed for UXO risk assessment. The first would apply to prioritizing UXO sites across the nation to determine which sites pose the greatest risks. The second process would be used for detailed evaluations of appropriate responses to UXO at specific sites. Within each process, separate methods are needed for assessing explosion risks and for assessing other constituent risks. None of the existing methods developed for UXO risk assessment that we reviewed is suitable for any of these applications, although elements of some of the methods could serve as input to new methods.

We recommend the following steps toward improving risk-based prioritization of UXO sites:

- **Develop a new UXO prioritization process that (1) sorts sites into bins by explosion risk and (2) within these bins, sorts sites by munitions constituents risks.** The suggested prioritization

process would preserve the information about the two separate risk types: although sites would be grouped first according to explosion risk, within these groups the sites would be ordered by munitions constituents risk. Policymakers could then decide how to distribute limited resources among sites with different combinations of explosion and constituent risk.

- **Develop a new process for sorting sites by explosion risk (stage one of the prioritization process).** The existing Risk Assessment Code could provide elements for the new process, but stakeholder concerns would need to be addressed.
- **Use the EPA Hazard Ranking System or Defense Environmental Restoration Program Relative Risk Site Evaluation Primer for sorting sites by munitions constituents risks (stage two of the prioritization process).** These methods are well established and well accepted. There is no need for a new approach for munitions constituents risk ranking, since the behavior of these contaminants and the risks they pose are analogous to those of chemical contaminants found at non-UXO hazardous waste sites.
- **Produce two UXO site priority lists: one for sites with known and documented future land use and another for sites with uncertain future land use.** Having two lists would prevent manipulation of the process by choosing the least restrictive land uses. Also, it would allow policymakers to decide how to trade off current and future risks when allocating funds. The lists could be updated periodically (e.g., annually) or as often as new information became available.
- **Appoint an independent technical review board and an advisory committee of stakeholders to oversee development of the prioritization process.** The technical board would consist of independent experts in risk assessment and explosive ordnance disposal. The advisory committee would include representatives of the different groups of stakeholders (state regulators, federal regulators, Native Americans, members of the public, military personnel) involved at UXO sites.

We recommend the following steps for improving risk-based selection of remedies at UXO sites:

Summary xxiii

- **Use available processes (RAGS and ERAGS) for site-specific assessment of munitions constituents risks.** RAGS and ERAGS are well established for assessing risks of chemicals in water and soil, and there is no need for the Army to develop a new method.
- **Develop a new, probabilistic approach using fault and event trees or similar methods for site-specific assessment of explosion risks.** None of the available UXO explosion risk assessment methods by itself satisfies technical criteria for an effective risk assessment method, so a new approach is needed. Many other agencies use probabilistic risk assessment tools to assess risks of acute events analogous to UXO explosion.
- **Create a set of fault and/or event trees at the national level that could serve as templates for local assessments and guidelines for use of those trees in computing probabilistic risk estimates.** One advantage of tree-based approaches is that they are easily adapted to local conditions, but having national models in place would allow for efficient development of trees at the local level. With significant modification and stakeholder input, some of the exposure scenarios developed for the Ordnance and Explosives Cost-Effectiveness Risk Tool might provide elements of UXO probabilistic risk assessments.
- **Involve an independent technical review board and an advisory committee of stakeholders from the beginning of development of the probabilistic site-specific risk assessment process.** Seeking input from independent reviewers as the risk assessment process is conceived will ensure that it is technically sound and that it meets the needs of stakeholders.

ACKNOWLEDGMENTS

Many people have contributed to this project, giving generously of their knowledge and time. We deeply appreciate their contributions. We specifically acknowledge the invaluable assistance provided by the following individuals:

David Anker, Center for Army Analysis

Roger Azar, CH2M Hill

Larry Barb, Army Corps of Engineers

David Bennett, Environmental Protection Agency

Heather Black, Malcolm Pirnie, Inc.

James Coppola, Army Corps of Engineers

David Douthat, Army Corps of Engineers

B. John Garrick, Garrick Consulting

Brian Helmlinger, Army, Office of the Deputy Chief of Staff for Operations

Scott Hill, Army Environmental Center

Juan Koponen, Army Corps of Engineers

J.C. King, Army, Office of the Deputy Assistant Secretary
(Environment, Safety, and Occupational Health)

Mary Ellen Maly, Army Environmental Center

James Manthey, Army Corps of Engineers

Jeff Marquesee, Office of the Secretary of Defense, Strategic Environmental Research and Development Program

John Michaud, Environmental Protection Agency

Joseph Murphy, Army Environmental Center

D. Warner North, NorthWorks, Inc.

Kathleen O'Halloran (retired), Army, Office of the Assistant Chief of Staff for Installation Management

John Selstrom, Office of the Deputy Under Secretary of Defense (Installations and Environment)

Ken Shuster, Environmental Protection Agency

Joe Skibinski, Science Applications International Corp.

Tom Symalla, Army Environmental Center

Connie Van Brocklin, Army, Office of the Assistant Chief of Staff for Installation Management

Jeff Waugh, Army Environmental Center

James Woolford, Environmental Protection Agency

Richard Wright, Army Environmental Policy Institute

Robert York, Army Environmental Center

Gail Youngblood, Army, Presidio of Monterey

We would also like to thank John Dumond, Director of the RAND Arroyo Center's Military Logistics Program, and Tom McNaugher, Director of the Arroyo Center, for their continuing support for our efforts. Special thanks are also due to Jerry Sollinger and Leigh Stallings for their valuable assistance in helping us prepare the manuscript.

ABBREVIATIONS

ATSDR	Agency for Toxic Substances and Disease Registry
BRAC	Base Realignment and Closure
CERCLA	Comprehensive Environmental Response, Compensation and Liability Act
CHF	Contaminant hazard factor
CSDP	Chemical Stockpile Disposal Program
DDESB	Department of Defense Explosives Safety Board
DERP	Defense Environmental Restoration Program
DNT	Dinitrotoluene
DoD	Department of Defense
DOE	Department of Energy
EMI	Electromagnetic induction
EPA	Environmental Protection Agency
ERAGS	Ecological Risk Assessment Guidance for Superfund
ERG	External Review Group
ERPS	Environmental Restoration Priority System
ES	Exposed site
ESHA	Explosive Safety Hazard Assessment

FAA	Federal Aviation Administration
FUDS	Formerly Utilized Defense Site
GAO	General Accounting Office
HMX	High-melting explosive
HRS	Hazard Ranking System
IR3M	Interim Range Rule Risk Methodology
MMR	Massachusetts Military Reservation
MPF	Migration Pathway Factor
NASA	National Aeronautics and Space Administration
NCRB	Natural and Cultural Resources Bank
NPL	National Priorities List
NRC	Nuclear Regulatory Commission
OE	Ordnance and explosives
OECert	Ordnance and Explosives Cost-Effectiveness Risk Tool
OERIA	Ordnance and Explosives Risk Impact Analysis
OSHA	Occupational Safety and Health Administration
PES	Potential explosive site
PRA	Probabilistic risk assessment
QRA	Quantitative Risk Assessment
QRAS	Quantitative Risk Assessment System (software)
RAC	Risk Assessment Code
RAGS	Risk Assessment Guidance for Superfund
RDX	Royal demolition explosive
RF	Receptor factor
ROC	Receiver operating characteristics
RRSE	Relative Risk Site Evaluation

Abbreviations xxix

SAFER	Safety Assessment for Explosive Risk
SAPHIRE	Systems Analysis Programs for Hands-on Integrated Reliability Evaluation
SDWA	Safe Drinking Water Act
TNT	Trinitrotoluene
TSPA	Total System Performance Assessment
UXO	Unexploded ordnance

Chapter One
INTRODUCTION

The presence of unexploded ordnance (UXO) on civilian land is a growing domestic concern due to the increasing number of closed military bases. Since the end of the Cold War, approximately 20 percent of the land owned by the Department of Defense (DoD) has been slated for transfer to civilian uses under the congressionally mandated Base Realignment and Closure (BRAC) program. UXO is present on some portion of this land, but precisely how much is not known. Chemicals, such as explosives, that are components of military munitions also may be present in soils and groundwater.

Until recently, civilian encounters with UXO were limited because of the restrictions on access to military property. But as bases close and access restrictions lift, there is concern that UXO risks will increase, unless remediation or preventive measures are taken. Congress has signaled its interest in this issue by enacting legislation that requires the DoD to develop an inventory of UXO sites, a protocol for establishing response priorities among them, and other tools to advance the cleanup and stewardship of these sites.

This report critically evaluates and recommends improvements to methods for assessing the risks from UXO and munitions constituents at domestic closed, transferred, and transferring military installations. It also examines methods for ranking risks among UXO sites for programmatic priority setting. The report was prepared at the request of the Army Assistant Chief of Staff for Installation Management. At the time this study was initiated, the Army had the lead responsibility on behalf of DoD for developing risk assessment processes for UXO sites.

DoD personnel have already developed risk assessment and prioritization methods for UXO sites. However, agreement on which, if any, methods are most appropriate is lacking among those involved at UXO sites. The DoD has been unsuccessful in the promulgation of a standard risk assessment protocol in part because of this lack of consensus. Technical complexities and uncertainties associated with UXO sites also have contributed to the difficulties of UXO site risk assessment. In addition, some stakeholders object to any use of risk assessment as an endorsement of nonzero risk. The purpose of this report is to help the Army and DoD evaluate the strengths and weaknesses, from a technical perspective, of existing UXO risk assessment methods (including those for detailed site evaluation and those for prioritization) and to assess whether methods used by other federal agencies might provide suitable models.

STUDY TASKS

The Army asked RAND Arroyo Center to carry out the following three tasks related to development of a consistent DoD strategy for assessing risks at UXO sites:

1. Conduct a preliminary analysis of ongoing efforts in UXO risk assessment. The analysis should include the (a) DoD Explosives Safety Board standards for explosives safety risk; (b) Defense Environmental Restoration Program *Relative Risk Site Evaluation Primer*; (c) Army Corps of Engineers' *Ordnance and Explosives Cost-Effectiveness Risk Tool*; (d) Army Environmental Center *Interim Range Rule Risk Methodology*; (e) Environmental Protection Agency (EPA) *Policy for Addressing Ordnance and Explosives at Closed, Transferring, and Transferred Ranges and Other Sites* and *Handbook on the Management of Ordnance and Explosives at Closed, Transferred, and Transferring Ranges*; and (f) any methods funded or developed by the Environmental Security Technology Certification Program and Strategic Environmental Research and Development Program.¹

¹A review of the new Munitions Response Site Prioritization Protocol being developed by the DoD is not included because this protocol was not available when this report was written.

2. Study methods used by the Department of Energy, National Aeronautics and Space Administration, and others to evaluate and measure risk of low-probability and high-consequence events such as the catastrophic risk of nuclear power plant failure or space shuttle explosion. Determine whether elements of these methods might be applicable to the assessment of risks at UXO sites.
3. Develop options and make recommendations on how the Army could develop a risk assessment/risk management protocol for UXO response actions that integrates UXO explosives safety and environmental contamination risks.

STUDY METHODS

This review is based on a thorough assessment of available technical literature, both published and unpublished, and on interviews with numerous individuals involved in UXO site risk assessment. To evaluate existing risk assessment methods for UXO sites, we reviewed documentation describing the methods; tested software where applicable; read existing critiques of the methods where available; and consulted with method developers, regulatory officials, Army field personnel, and other stakeholders involved at UXO sites. To assess whether the risk assessment methods of other agencies might be useful for UXO sites, we identified selected agencies with risk management problems that bear some similarity to UXO risks. Then we assessed in detail the methods that appear to be most applicable.

Underpinning our evaluation of existing UXO risk assessment methods is a set of characteristics that describe an “ideal” UXO risk assessment process. We derived these characteristics from literature on risk assessment, consultations with others knowledgeable about risk assessment, and our own expertise. We compared each existing UXO risk assessment method to the set of evaluation criteria to highlight strengths and weaknesses of the methods and to help identify gaps in the existing tool set. We also consulted literature concerning the management of disparate risks.

HOW THIS REPORT IS ORGANIZED

The remainder of this chapter provides background information about UXO sites, including information about the number of sites, risks at those sites, and difficulties in detecting and clearing UXO. Chapter Two presents background information about risk assessment and general principles for a credible risk assessment method. Chapter Three reviews UXO risk assessment and prioritization methods that were available as of approximately June 2002. Chapter Four describes features of risk assessment methods used by other agencies that could be useful for UXO risk assessment in the future. Chapter Five recommends possible strategies for developing standard processes for UXO risk assessment.

CATEGORIES OF UXO SITES

UXO sites generally are grouped into two categories: those that were or are being transferred under BRAC, and those that were closed prior to BRAC. Sites in this second category are known as Formerly Utilized Defense Sites (FUDS). An example of a FUDS site with UXO is the Spring Valley area of Washington, D.C. Here, chemical weapons were disposed of after World War I and were not discovered until a contractor digging a utility trench in 1993 uncovered a disposal pit containing 141 UXO items (Jaffe, 2003; Nielson and Anderson-Hudgins, 2002).

NUMBER OF UXO SITES

Estimates of the total number of UXO sites vary substantially. The Defense Science Board Task Force on UXO estimated that there are 1,500 domestic UXO sites (Department of Defense, 1998). EPA has estimated that more than 7,500 sites may have UXO.² The DoD's most recent assessment, known as the "Advance Range Survey," estimated that there are 859 military installations and properties with closed, transferred, and transferring ranges, but each property typically contains multiple UXO sites (Maly, 2002). In response to con-

²From unpublished information provided by Ken Shuster of the Environmental Protection Agency, August 28, 2000.

gressional legislation requiring establishment of a permanent inventory of UXO sites, the DoD is planning to conduct a comprehensive inventory of UXO sites by December 2003 (Maly, 2002).³ Once this inventory is completed, more detailed information about the extent of the UXO problem will be available. However, uncertainties may remain due to the possible existence of former UXO sites that were closed long ago and may have been missed during the inventory.

SOURCES AND TYPES OF UXO

The presence of UXO is inevitable on any land that the military used for training or weapons development and testing. No type of munition explodes 100 percent of the time when fired. Dud rates are as high as 10 percent, depending on the type of munition, according to the Defense Science Board Task Force on UXO (Department of Defense, 1998). Surveys in Laos and Cambodia after the Vietnam War indicated that 10 to 30 percent of bombs dropped on these countries failed to detonate (Lauritzen, 2001). When live munitions are fired on a range over a period of decades, a large amount of UXO can accumulate. Even if the military periodically clears most of the surface UXO, ordnance that is buried beneath the soil due to the force of the initial impact or to weathering will remain behind.

The types of munitions found at domestic UXO sites vary widely depending on the types of military activities that took place at the site. UXO can range from small-arms ammunition to bombs weighing up to a ton. Other types of munitions include artillery rounds, mortars, aircraft cannon, tank-fired projectiles, submunitions (which are designed to scatter over a large area), rockets, guided missiles, grenades, torpedoes, mines, chemical munitions, bulk explosives, and pyrotechnics (see Figure 1.1 for an example). Each of these munitions types differs in the amount of explosive contained, the depth in the ground to which it is likely to penetrate, the sensitivity of the unexploded item to detonation, and the potential for explosives and other contaminants to leak into surrounding soil.

³Congress spelled out requirements pertaining to UXO in the National Defense Authorization Act of 2002, sections 311–313.



PHOTO: Supplied by Fort McClellan.

Figure 1.1—60mm HE Mortar Found at Fort McClellan

RISKS FROM UXO SITES

Two kinds of risks are associated with UXO sites. The first is the obvious risk of explosion. Human exposure to exploding UXO causes serious injury, dismemberment, or death. In this report, we call such risk “explosion risk.” The second type of risk is due to the infiltration of chemicals from the munitions (primarily explosives) into soil and groundwater; we call this second type of risk “munitions constituents risk.”

A number of injuries and fatalities have occurred at domestic military ranges over the years due to accidental UXO explosions. The Arroyo Center was unable to obtain complete information about all UXO incidents as part of this project but did receive isolated reports. According to a report from the General Accounting Office (GAO), 24

incidents involving UXO were documented in an EPA survey of records from 61 DoD facilities containing 203 inactive, closed, transferred, or transferring ranges (the report does not indicate the time span over which the incidents occurred) (GAO, 2001). The incidents led to three deaths and two injuries. A separate incident occurred in San Diego in 1983, on a property that was once part of Camp Elliott. Two young boys were killed when they found an unexploded bomb in an open gully near their backyards. In March 2001, homeowners in the same San Diego community discovered an armed bomb when digging for a backyard pool (North County Times, 2001), but fortunately there were no injuries associated with this incident. As a final example, according to published newspaper reports and an unpublished search of court records gathered from the Fort Ord area, there were five incidents in which children were killed or seriously injured by UXO at Fort Ord between 1943 and the base's closure in 1994 (Monterey Peninsula Herald, 1943, 1949, 1976). Since the base closed, there have been many incidents of trespassing on former range areas, including one in which schoolchildren loaded their backpacks with training grenades. Incidents of trespassing have slowed since Fort Ord installed a concertina-wire fence around the main range and instituted a program to educate local schoolchildren about risks from UXO.

The second type of risk that may be present at UXO sites is due to munitions constituents (primarily explosives, though lead may be a concern at small-arms ranges) in soil and groundwater. Sources of munitions constituents include explosives residues left after munitions have detonated either partially or completely, residues from blow-in-place operations to destroy duds, open burning of excess explosives, corrosion of UXO items, and breakage of munitions without detonation (Jenkins, Hewitt, et al., 2002). The most common explosives are trinitrotoluene (TNT), royal demolition explosive (RDX), and high-melting explosive (HMX); various isomers of dinitrotoluene (DNT) are also often present as byproducts of TNT degradation. Concerns about the presence of munitions constituents, in particular explosives, have increased since RDX was discovered in groundwater beneath the impact range at the Massachusetts Military Reservation (MMR) on Cape Cod. As a result of this contamination, the EPA in January 2000 ordered a complete halt to all training using live or training munitions at MMR.

A number of studies are under way to characterize the extent of soil and groundwater contamination at individual ranges (Jenkins, Hewitt, et al., 2002; Packer and Thomsen, 2002). For example, Jenkins, Hewitt, et al. sampled soil from ranges at six military installations and found various combinations of explosives present on portions of all the ranges. The concentrations of explosives varied with the types of ordnance fired (e.g., artillery, grenade, anti-tank), as well as with other factors such as environmental conditions and density of ordnance items. Explosives concentrations were generally low and attenuated rapidly with distance from the target areas (in both the horizontal and vertical directions), but most ranges had localized sources with high concentrations. For example, at a former anti-tank range at Fort Ord, the mean concentration of HMX in one of the target areas was 295,000 micrograms per kilogram of soil (micrograms/kg), but 15 meters from the target the mean concentration decreased to 1,440 micrograms/kg. The maximum concentration of HMX detected at this range was 587,000 micrograms/kg. At an artillery range at Fort Lewis, the concentration of TNT was generally less than 100 micrograms/kg, but in an area where a 155mm round had failed to completely detonate, the TNT concentration was 15,100,000 micrograms/kg.

Although studies have characterized explosive residues in soil and groundwater at isolated ranges, no comprehensive survey is available to indicate the extent to which explosives and other munitions constituents are present in soil and groundwater at all closed, transferred, and transferring ranges. The DoD is planning to collect this type of data as part of its ongoing initiative to develop a comprehensive inventory of UXO sites.

The health risks of exposure to low levels of explosives in soil and groundwater are highly uncertain and would need to be evaluated on a case-by-case basis. Exposure to TNT at high concentrations, such as among workers at ammunition plants, has caused blood disorders such as anemia and abnormal liver function, as well as allergic reactions to the skin and cataracts (ATSDR, 1995b). However, as for most contaminants, the effects of chronic exposures to low contaminant concentrations are highly uncertain. Based on animal studies, the level of TNT at which no adverse effects can be expected for long-term exposure is 1 mg/kg/day (ATSDR, 1995b). At high exposures, RDX can cause seizures in humans, but the effects of long-term, low-

level exposure are unknown (ATSDR, 1995a). Based on animal studies, the level at which no adverse effects are expected is 10 mg/kg day for chronic exposure (ATSDR, 1995a). No information is available about the human health effects of HMX, but animal studies indicate that it may be harmful to the liver or central nervous systems at doses above 100 mg/kg/day (ATSDR, 1997). Animal and human studies indicate that 2,4-DNT may cause fertility problems and nervous system disorders, as well as liver and kidney damage and a reduction in red blood cell count at doses above 0.1 mg/kg/day (ATSDR, 1998). Whether these explosives cause cancer is unknown, because animal studies are inconclusive, and no data are available from human studies (ATSDR, 1995a,b, 1997, 1998).

As an example of the potential variability of human health risks from munitions constituents, consider the case of the Fort Lewis range sampled by Jenkins, Hewitt, et al. At the maximally contaminated part of this range, a 50-kg individual would exceed the recommended health-based level of TNT (1 mg/kg/day) by consuming 0.003 kg per day of soil. However, at most of the locations across the Fort Lewis range, the same 50-kg individual would need to ingest 500 kg of soil per day for a long time to exceed the recommended health-based level of TNT.

UXO DETECTION AND REMOVAL PROCESS

The tools available for UXO detection and removal are essentially the same as those that the military has employed for clearing mines and explosive ordnance since World War II. Typically, the UXO clearance crew is equipped with a metal detector and a shovel. The crew first clears vegetation from the UXO area (using mechanical methods or controlled burning). Then, the team divides the area into grids, splits the grids into lanes (usually one meter wide), and slowly advances down each lane, swinging the metal detector close to the ground (see the example from Fort Ord in Figure 1.2). When the detector beeps, the crewmember either plants a flag to indicate that excavation will need to occur or starts digging with a shovel until the metal object is located. If the object is not a UXO item, it is excavated and laid aside. If it is a UXO item, it is either blown in place or carefully removed for later detonation.



PHOTO: Jackie MacDonald, RAND.

The ropes mark search lanes. The light-colored object in the foreground is a metal detector. The workman is digging in an area where the detector indicated that metal is present.

Figure 1.2—Searching for UXO at Fort Ord

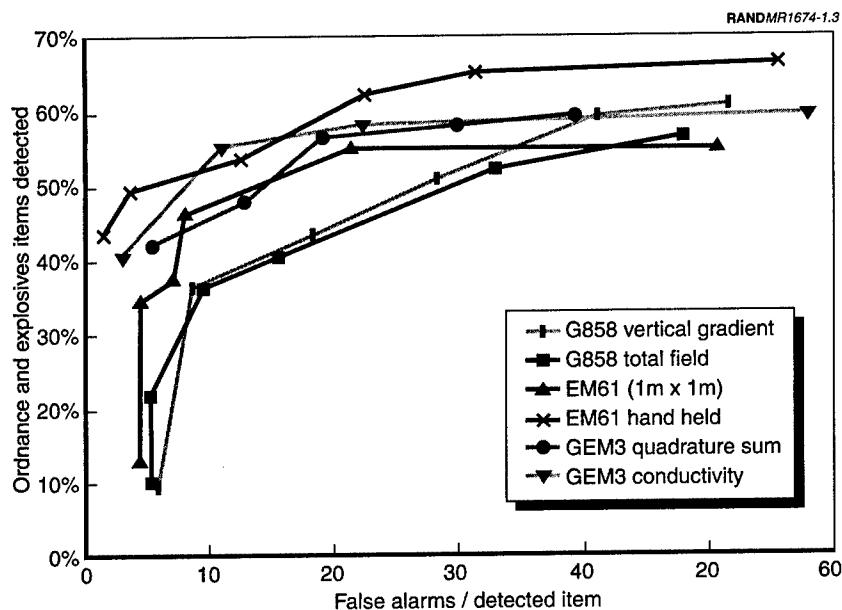
The metal detectors used for this process are either magnetometers or electromagnetic induction (EMI) systems. They have changed little in principle since World War II (although some systems are now equipped with mapping devices that store information about anomaly locations in a geographic information system). Magnetometers measure distortions in the earth's magnetic field caused by the presence of metal objects. EMI systems generate a magnetic field in the ground that induces current to flow in buried metal; this current in turn induces a secondary magnetic field with a voltage that is detected by the EMI instrument.

The overwhelming limitation of mine detection using magnetometers and EMI systems is the inability to discriminate UXO from metal clutter. The detectors are highly sensitive to small metal fragments, including shrapnel, bottle caps, bullet casings, soup cans, and other man-made clutter as well as natural metal in rock. The operator must therefore strike a balance between tuning the detector so finely that it generates an overwhelming number of false positive signals and not tuning it finely enough, in which case it misses too many UXO items. The balance between these two competing objectives is quantified by what is known as a “receiver operating characteristics” (ROC) curve. A ROC curve plots the probability of finding a buried UXO item (known as the “probability of detection,” or P_d) against the probability that a detected item will be a false alarm (known as the “probability of false alarm,” or P_{fa}). Both probabilities are plotted as a function of the threshold used to decide whether or not to make a declaration (e.g., the loudness of the tone produced by a magnetometer), thus defining a curve.

ROC curves vary not only with the detector but also with the location where the detector is employed and the radius of the area searched in response to an anomaly detection. Figures 1.3 and 1.4 show example ROC curves from field tests of EMI systems at Fort Ord.

As shown, in these field tests the number of false alarms per UXO item varied with the detector chosen, the search radius, and the desired probability of detection. Data from Fort Ord indicate that in practice, approximately 99 false alarms are excavated for every UXO item found: a total of 5.5 million items have been excavated at Fort Ord, but only 49,000 of them were UXO (Kevin Siemann, Fort Ord, personal communication, April 2002).

Compounding the false alarm problem is that no matter how careful the detector operator, the systems still miss UXO items. As shown in Figure 1.3, when the search radius for the Fort Ord study was 1.6 feet, the best-performing of the detectors evaluated in this test found fewer than 70 percent of UXO items, even when the threshold was finely tuned to generate a large number of false alarms. When the search radius was increased to 3.3 feet (meaning a hole of radius 3.3 feet was dug around every signal), the probability of detection increased to more than 95 percent, as shown in Figure 1.4. However,



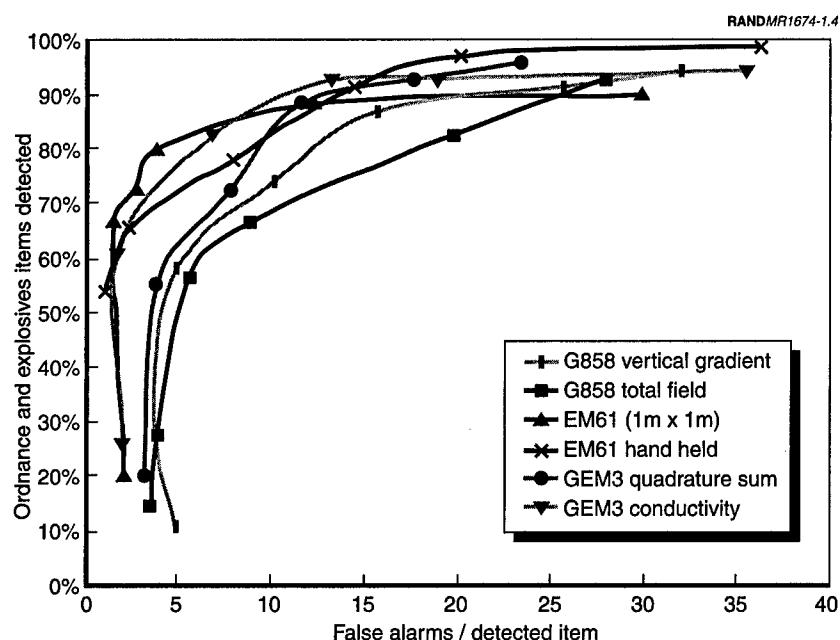
SOURCE: Army Corps of Engineers (2002).

The data for this curve were obtained using a search radius of 1.6 feet. The different curves show performance of the six detectors tested.

Figure 1.3—Example ROC Curve from Field Test of EMI Systems at Fort Ord (1.6-foot search radius)

even with a detection probability of 95 percent, UXO items will be left behind.

As an alternative strategy to ensure a greater likelihood that all UXO will be removed, some environmental regulators have proposed excavating entire UXO sites one foot at a time—essentially sifting the entire site. In this proposed process, the site would be cleared of surface ordnance. Then, it would be scanned once, and all anomalies would be excavated down to one foot of depth. The first foot of soil over the entire site would then be removed. The excavated soil would be sifted through a sieve that would isolate any UXO not detected in the first scan. The bottom of the excavated area would then be scanned with the detector, anomalies would be removed to an



SOURCE: Army Corps of Engineers (2002).

The data for this curve were obtained using a search radius of 3.3 feet. As in Figure 1.3, the six curves correspond to the performance of six different detectors in field tests.

Figure 1.4—Example ROC Curve from Field Test of EMI Systems at Fort Ord (3.3-foot search radius)

additional foot, the next foot of soil would be entirely removed, and the excavated soil would be sifted and set aside. This process would continue until no additional items were found. Figure 1.5 illustrates this proposed process.

The sifting approach has two drawbacks. First, as will be discussed in Chapter Two, sifting all UXO sites would exceed the historical level of the DoD environmental remediation budget, leaving few or no resources for other required environmental activities. Second, in many cases, UXO sites are located in areas with threatened and endangered species. Sifting permanently destroys the vegetation and soil and may be at odds with habitat conservation and species protection

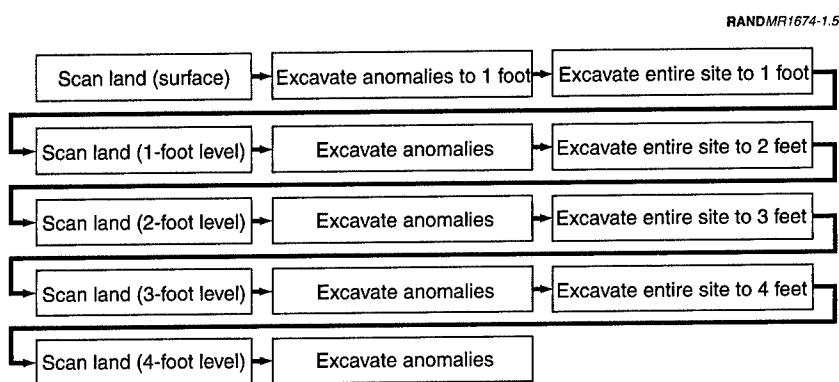


Figure 1.5—Example of a Process Proposed by State Regulators for Cleanup of a UXO Site

plans under the Endangered Species Act and comparable state laws. Ironically, the presence of UXO has prevented human intrusion at many locations and allowed species to survive where they would otherwise have perished. As an example, Fort Ord contains approximately 85 percent of the world's remaining rare and endangered plants found in a type of habitat known as maritime chaparral (Presidio of Monterey Directorate of Environmental and Natural Resources, 2001). Almost all of California's maritime chaparral outside Fort Ord has been destroyed by human development.

Several other UXO clearance protocols have been considered, including

- clearing only surface UXO;
- digging where a detector signals an anomaly, but only to a specified depth (e.g., two feet);
- digging wherever the detector signals until an anomaly is found;
- digging wherever the detector signals until an anomaly is found, then scanning the bottom of the resulting hole and digging again if the detector signals that another anomaly may be present;
- repeating the scan-and-dig approach two or more times; and
- sifting the entire site to various depths, as described above.

At most UXO sites, stakeholders have been unable to agree on which of these approaches is best. The underlying problem is that no approach short of sifting can guarantee that all UXO has been removed because of the limitations of metal detectors for scanning the subsurface. Risk assessment could be useful for evaluating the different degrees to which each alternative short of sifting reduces people's risk of death, physical injury, or illness from UXO and munitions constituents.

EFFORTS TO STANDARDIZE UXO RISK ASSESSMENT PROCESSES

The DoD has long recognized the potential value in having a standard process for assessing risks of UXO sites. In the absence of such a process, a new risk assessment approach must be developed for every site. Negotiating with regulators and other stakeholders about the legitimacy of the chosen process and whether the proposed clearance approach will sufficiently protect the public is a time-consuming, costly process. A standard risk assessment process that is widely endorsed by regulators and stakeholders would reduce the time and money spent developing a new approach for each site. Further, with direction from Congress, the DoD is required to develop a method of setting priorities among sites based on, among other considerations, the relative risks of UXO and munitions constituents.

The Army, on behalf of the DoD, has made numerous attempts to standardize UXO risk assessment. The most recent effort was the development of the Interim Range Rule Risk Methodology (IR3M). IR3M was intended to integrate the explosives safety and toxicological evaluations of different possible approaches for clearing UXO. Development of IR3M was part of a larger process to establish a uniform national approach—known as the Range Rule—for managing UXO sites. The Army spent two years, from 1995 until 1997, developing the draft Range Rule, which was published in the *Federal Register* (Department of Defense, 1997a). For the next three years, the Army negotiated with the EPA and other involved federal agencies (those intended to be receivers of DoD land once it is cleared of UXO) about necessary revisions. However, the agencies could not agree on several key process issues, the most important being which agency should have final authority to approve the UXO clearance approach

(MacDonald, 2001). As a result, DoD withdrew the Range Rule from consideration in November 2000. Work on IR3M halted at the same time. By that point, IR3M had been fully developed and had undergone limited testing and evaluation.

The IR3M was not the first attempt to standardize the UXO risk assessment process. Previously, the Army Corps of Engineers had developed risk assessment tools to be used in managing UXO sites under the FUDS program. These methods included the Ordnance and Explosives Cost-Effectiveness Risk Tool (OECert), the Ordnance and Explosives Risk Impact Analysis (OERIA), and the Risk Assessment Code (RAC). The latter method was intended to indicate relative risks of UXO explosions among sites. EPA objected to all of these methods on various grounds.

In the absence of widely accepted UXO risk assessment methods, prioritization of UXO sites using risk criteria—as is now required by Congress—is not possible. Further, selection of a clearance process for UXO sites is a cumbersome, inefficient process. DoD personnel charged with overseeing UXO clearance activities must first decide whether to try using one of the available risk assessment methods—IR3M, OECert, OERIA, or RAC—or to develop a new method specific to the site. In some cases, a hybrid of an existing method is chosen. For example, hybrids of IR3M are being used at Fort Ord and at Adak Island, Alaska. DoD site managers then must convince regulators and the public that the chosen approach is technically sound. At this stage, disagreements about risk assessment often slow efforts to remove UXO or halt them altogether, which exacerbates the risks by increasing the potential for human exposure.

This report first assesses the previous efforts to develop consistent UXO risk assessment methods for two separate purposes:

1. Risk-based prioritization for deciding how to allocate funds and other resources at the national level.
2. Site-specific evaluation to help make decisions about the UXO clearance process and the suitability of the site for alternative land uses.

Then, it reviews how other agencies have handled similar risk problems. Finally, it recommends a path forward for developing standard

Introduction 17

UXO risk assessment processes and methods. Throughout, we distinguish between risk assessment for prioritization versus for site-specific evaluation and between risk assessment for explosion risks versus for munitions constituents risks.

Chapter Two

POTENTIAL FOR A RISK-BASED APPROACH TO UXO RESPONSE

In the absence of sifting each UXO site to a depth of at least several feet, complete elimination of UXO can never be guaranteed. This is because technologies for detecting UXO are imperfect: no existing technology can provide complete assurance that every buried UXO item has been located and removed. Sifting all the soil may be feasible in a limited number of cases, but usually it will not be possible because of either cost constraints or concerns about irreparable damage to ecosystems and threatened and endangered species. Each alternative other than sifting will leave some level of residual risk. A credible and technically sound risk assessment process is essential for evaluating these residual risks and for choosing the best among imperfect choices for UXO detection and clearance.

This chapter first explains the basis for and history of risk assessment. It then describes how risk assessment could be used to improve management of UXO sites and explains why risk assessment has not yet played a major role at UXO sites. The final section establishes a set of criteria that a credible UXO risk assessment method should meet.

The problem of assessing risks from UXO sites is not unique. Throughout history, humans have had to make difficult choices about managing risks in the face of technical, information, and resource limitations. Risk assessment has evolved as a discipline to meet this need. There is a wealth of knowledge about risk assessment to draw from in informing choices about UXO sites. Although the use of risk assessment at UXO sites has been questioned, we be-

lieve it is a necessary process in public approaches to evaluating alternative options for UXO response.

WHAT IS RISK ASSESSMENT?

A risk assessment is a systematic process for identifying potential hazards and the likelihood that those hazards will cause harm. In a study that has served as a guide for the conduct of risk assessments in the federal government, the National Research Council of the National Academy of Sciences defined risk assessment as “the use of the factual base to define the health effects of exposure of individuals or populations to hazardous materials and situations” (National Research Council, 1983). Risk assessment involves more than producing a single number or other parameter to describe risk; it involves an organized process for characterizing the risk in question (Graham, 1995).

In general, human health risk assessment as defined by the National Research Council consists of some or all of the following four steps:

1. **Identify the hazard.** Determine whether a particular contaminant is or is not causally linked to health effects.
2. **Quantify the dose-response relationship.** Determine the relationship between the amount of contaminant to which an individual or population is exposed and the probability that adverse health effects will occur.
3. **Assess the exposure.** Determine the extent to which humans will be exposed to the contaminant, either before or after regulatory controls are implemented.
4. **Characterize the risk.** Describe the nature and magnitude of the risk to humans, including the uncertainty associated with the analysis.

The details of how these steps are carried out will vary with the nature of the risks involved and the information available to analyze them. Not all of the steps are necessary for every risk problem. For example, the first two steps are trivial in the assessment of risks from the potential explosion of a UXO item: if the ordnance explodes in the presence of a person, either serious injury or death is a certain

outcome. On the other hand, for assessing risks of munitions constituents in soil and groundwater, the first two steps would require a careful analysis of available information about the health effects of ingestion of low doses of explosives such as TNT and RDX.

HISTORY OF RISK ASSESSMENT

A common misconception about risk assessment is that it is a product of the late 20th century, created to address new concerns about technological hazards and environmental contaminants. In fact, humans have employed risk assessment to make difficult decisions literally from the earliest organized civilizations.

As early as 3200 B.C., the Babylonians consulted a priest-like sect known as the Asipu to analyze alternative actions for coping with risky situations (Covello and Mumpower, 1985). Later, in about 1800 B.C., Hammurabi, King of Babylon, formalized the concept of buying insurance to protect against risks. The earliest insurance policies covered losses of cargo at sea in exchange for payment of interest to moneylenders (Bernstein, 1996). Life insurance was another early form of risk management, instituted by Roman *collegia* (Bernstein, 1996). Concern about contaminants in the environment also dates to ancient times. The Greeks and Romans recognized the toxicity of lead, mercury, and fumes from burning charcoal (Graham, 1995).

The effort to quantify the likelihood that a risky event will occur dates to the development of probability theory nearly 400 years ago. The earliest use of probability for evaluating risks was in the establishment of life insurance premiums, in which insurance adjustors would determine the minimum premium necessary to cover the costs of a death benefit (Bernstein, 1996). In the seventeenth century, the French monk Antoine Arnauld observed, "Fear of harm ought to be proportional not merely to the gravity of the harm, but also to the probability of the event" (Bernstein, 1996).

Great strides in the ability to conduct quantitative assessments of risks to humans from environmental contaminants have occurred since about 1960 in response to two developments: (1) nuclear power and (2) concerns about the health effects of man-made chemicals. The former development led to advances in methods for determining the likelihood of acute, undesired events such as system

failures followed by catastrophic releases of contaminants. The latter led to advances in methods for estimating the likelihood and consequences of human exposures to dispersed contaminants in the environment.

Concerns about the potential for accidents at nuclear power plants led to the development of probabilistic risk assessment—a set of processes and tools for predicting the probability that a certain undesired event could occur. Congress passed the Atomic Energy Act of 1954 to encourage civilian uses of nuclear power. Congress empowered the Atomic Energy Commission (later reorganized into the independent Nuclear Regulatory Commission, NRC) to establish a regulatory program to assure that public health and safety would be protected. The technical foundation for the safety regulations was established in the early days and remains firmly established today. The regulations are based on traditional industrial codes and standards supplemented by detailed specialized requirements tailored to the specific issues of nuclear power plant safety.

Without risk assessment methods, there is no way to know either the probability of large accidents or their consequences, given a specific design. The need for such predictions necessitated the development of methods to assess both the likelihood and the consequences of plant failures and releases to the environment. Analysts adapted methods that had evolved for analyzing the reliability of engineered systems, especially missiles, to the problem of nuclear power plant risk assessment. For example, to help identify potential triggers of plant failure and radioactive release, they used approaches known as “fault-tree” and “event-tree” analysis, which were first developed by Bell Laboratories to improve the reliability of Minuteman missiles (Haimes, 1998). The first comprehensive analysis of nuclear power plant risks, known as the Reactor Safety Study (or Report WASH-1400), was sponsored by the NRC and completed in 1975 (NRC, 1975; see also Lewis et al., 1978). Although this report and its methods were criticized by some members of the public and the NRC itself was reluctant at first to embrace the methods as a supplement to their traditional approaches to understanding reactor safety, after the accident at Three Mile Island in 1979 the NRC began to rely on probabilistic risk assessment to evaluate the effectiveness of its regulatory programs. In recent years, the NRC has gone further and be-

gun to use these methods to evaluate plant-specific safety issues or design-change proposals.

At the same time as quantitative methods were being developed for assessing risks from nuclear power plants, significant advances occurred in the capability to assess human health risks from releases of contaminants into the environment. The demand for such assessments evolved in response to several events. Atmospheric testing of nuclear weapons—especially a 1954 blast that contaminated 43 residents of the Marshall Islands and 14 Japanese fishermen—led to calls to stop atmospheric testing and concerns about the transport of radioactive contaminants in the environment (Rechard, 1999). Rachel Carson's 1962 book *Silent Spring*, which documented the dangers of widespread pesticide use, reflected the growing concerns about dispersal of man-made organic chemicals in the environment (Carson, 2002). Releases of hazardous contaminants at Love Canal in New York and Times Beach in Missouri led to evacuation of residents in both locations due to concerns about exposure to man-made organic chemicals. By 1976, the EPA published its first guidelines on assessing risks of cancer from chemicals in water, air, and soil. During the 1980s and 1990s, the EPA developed approximately two dozen guidance documents (more than 5,000 pages) on conducting risk assessments of environmental contaminants (Paustenbach, 1995).

The risk assessment process is institutionalized in a number of U.S. government programs, not just for assessing risks of nuclear power plants and contaminants in the environment, but also for a wide range of other applications. Examples include food safety regulation by the Food and Drug Administration, safety management in the space program by the National Aeronautics and Space Administration (NASA), and regulation of civil aviation safety by the Federal Aviation Administration (FAA). Stakeholders unhappy with the outcome of risk assessments have challenged the risk assessment process in court, but in a 1980 case concerning occupational exposure to benzene, the Supreme Court upheld the use of risk assessment as a viable method for establishing regulatory standards (Graham, 1995). The Supreme Court decision set a precedent, and challenges to the legitimacy of risk assessment have been overturned in a number of court cases since then (Graham, 1995). Chapter Four describes uses of risk assessment in the government in more detail and illustrates

how some of the lessons from other federal risk assessments might apply to UXO sites.

POTENTIAL APPLICATIONS OF RISK ASSESSMENT TO UXO PROBLEMS

Risk assessment has two broad applications for the UXO problem: (1) programmatic prioritization and (2) site-specific assessment. The first application is to help optimize the allocation of limited national budgetary resources available for UXO response. Congress in the National Defense Authorization Act of 2002 mandated that the DoD develop a prioritized list of UXO sites and consider risk in establishing priorities. The second application includes assessing baseline risks at a site, without any UXO clearance, and determining how different approaches to UXO clearance will reduce the risk. Table 2.1 summarizes the types of questions that risk assessments for these two purposes can help to answer.

CHALLENGES OF USING RISK ASSESSMENT FOR UXO SITES

Two major impediments have prevented the use of risk assessment as an input to decisionmaking about UXO sites. First, representatives of the public, the EPA, and some state and local agencies have contended that the use of risk assessment implies that some level of risk above zero is acceptable. These stakeholders would like all risk from UXO eliminated. In a memo written to RAND on August 26, 2002, James Woolford, the director of EPA's Federal Facilities Restoration and Reuse Office, observed, "regulators and stakeholders do not generally consider any level of UXO risk to be acceptable, making any quantified risk higher than zero a hard sell indeed." Second, most environmental regulators are unfamiliar with the type of risk posed by UXO. Regulators are accustomed to considering long-term risks from exposures to low levels of contaminants in the environment, but not the possibility of immediate injury or death. Risk assessment is a required step in the regulatory decisionmaking process at hazardous waste sites, but the types of risk assessment methods used at these sites do not apply to UXO because of the nature of the risks involved.

Table 2.1
Questions Addressed by the Two Applications of Risk Assessment

Risk Assessment for Establishing Program Priorities	Risk Assessment for Detailed Evaluation of Sites
Of all installations in the United States or in a specific state, which poses the highest immediate risk to the public?	At an individual installation (for example, Fort Ord), what is the likelihood that the surrounding population will experience an adverse health effect from UXO at a specific location on that installation?
Which installation poses the highest risk of injury to individuals or populations due to UXO detonation?	To what extent would surface clearance of UXO reduce the risks of such adverse effects occurring?
Which installation poses the highest risk to populations due to munitions constituents in the environment?	To what extent would clearance to a specific depth reduce risks from UXO?
What is the relative difference in risk between installation A and installation B?	How will future land use changes affect the residual risk? Is it possible to reduce risk enough to render an area safe for new uses?
	Which areas of the installation pose the highest risks?
	What is the relative magnitude of the risk from munitions constituents as compared to the explosion risk?

Although eliminating all risk from UXO would be ideal, in practice this is not possible. To eliminate all risk, every UXO item would have to be unearthed. The only way to guarantee this is to sift the entire site down to the maximum penetration depth of the UXO encountered at that particular site. In some cases, this depth may be two feet; in other cases it may be ten feet or more. The costs of shoveling and sifting vast amounts of earth are astronomical. For example, a preliminary analysis of data from an actual closed range estimated that the cost of excavating to a depth of four feet (the maximum depth of UXO at that range) would be approximately \$125,000 per acre. In contrast, the analysis showed that if each anomaly found by a metal detector were excavated, the cost would be about \$10,000 per acre (Mendez, 2002). (Actual costs for anomaly excavation from this

site have proven to be about \$20,000 per acre.) The total budget for all environmental remediation programs in the DoD in 2001 was \$2.1 billion (Department of Defense, 2002). At a cost of \$125,000 per acre for sifting UXO sites, the entire DoD environmental restoration budget from 2001, which covers all environmental activities at all active as well as closed installations, would allow for sifting a total of 16,800 acres. The DoD has estimated that over 16 million acres of land on closed, transferred, and transferring ranges are potentially contaminated with UXO (GAO, 2001)¹. In the absence of budget resources being able to eliminate all UXO risk, the DoD needs a credible method to assess the extent to which alternatives that are short of sifting but are financially and technically feasible have the potential to reduce risks.

An additional impediment to using risk assessment as an input to decisionmaking at UXO sites is the nature of the UXO hazard. The EPA has extensive manuals providing methods for assessing the risks of contaminants that have dispersed in air, water, and soil at relatively dilute concentrations. However, no such guidance is available for quantifying the explosion risk from UXO. Existing EPA methods cannot be used for UXO explosion risks because these latter risks are qualitatively different from risks of contaminants that have been diluted in water, soil, or air. Table 2.2 summarizes important differences between explosion risks and risks from contaminants such as solvents or metals that have dispersed in soil and groundwater (the usual types of contaminants of concern at hazardous waste sites). The table illustrates these differences according to the first three stages of a health risk assessment as defined by the National Research Council. The final stage, risk characterization, should be similar for both types of risks.

The first stage of risk assessment is the hazard identification. That is, the first question the risk assessor must answer is whether the contaminant in question is known to cause the adverse health effects of concern. For some contaminants in soil and groundwater, answering this question poses a challenge. TNT is a good example. Although the EPA has categorized TNT as a possible carcinogen, data are not available through either the EPA or the Agency for Toxic Sub-

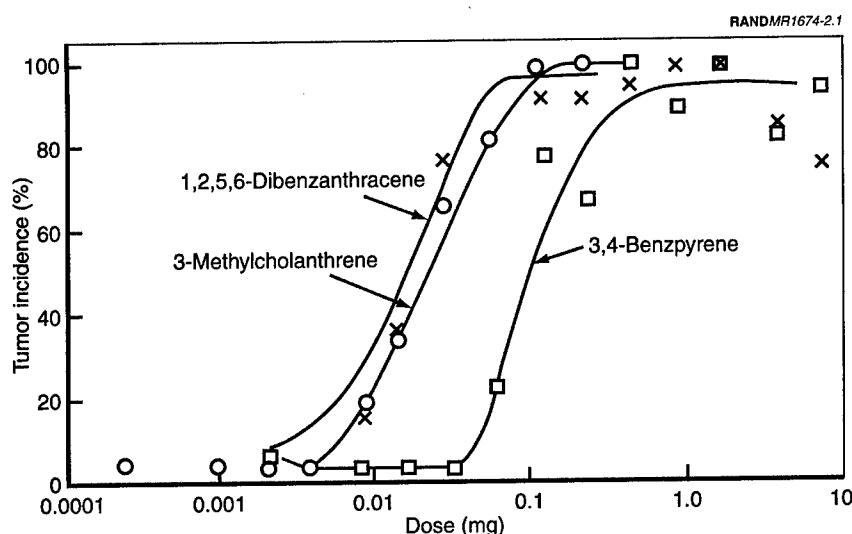
¹DoD was in the process of updating this estimate at the time we prepared this report.

Table 2.2
Differences Between Explosion Risks and Munitions Constituents Risks

Element of Risk Assessment	Explosion Risk Characteristics	Munitions Constituents Risk Characteristics
Hazard identification	Acute, short-term	Chronic, long-term
Dose-response assessment	Irrelevant: exposure does or does not result in injury or fatality, with no relation to "dose" (although the severity of injury may be related to the magnitude of the explosion)	Probability of adverse health effect generally increases with dose
Exposure assessment	Localized	Dispersed

stances and Disease Registry (the usual sources for such data) to confirm the cancer potency of TNT. On the other hand, noncancer health effects (such as anemia, aplastic anemia, and hepatitis) are well documented (ATSDR, 1995b). In contrast, for assessing the UXO explosion risk, the answer to the question of whether an explosion causes health problems is evident.

The second stage of risk assessment is the dose-response assessment. In this stage, the risk assessor must establish the relationship between the dose of the contaminant to which an individual might be exposed and the onset of adverse health effects. This is accomplished with the use of a "dose-response" curve, which shows the probability that an individual will contract a specific illness upon ingesting or coming in contact with a given contaminant concentration (see Figure 2.1 for an example). In most cases, there is a dose below which no adverse effects are expected to occur; above that dose, the likelihood of becoming ill increases with increasing dose, as in the examples in Figure 2.1. For assessing UXO explosion risks, the concept of a dose-response curve is irrelevant. The relationship between exposure to UXO and adverse health effects is binary: either the UXO does not explode and no harm is caused, or the UXO explodes and the person is seriously injured or killed. The severity of the injury may be related to the amount of explosive present, but in general no



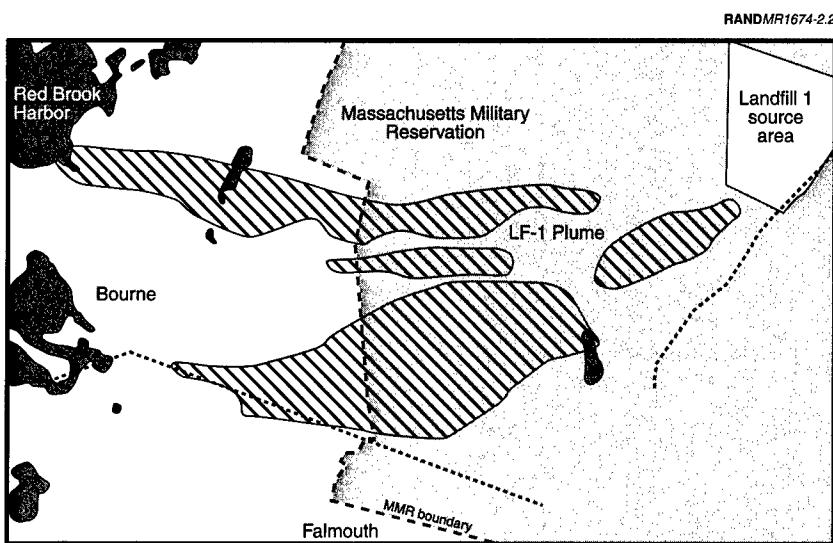
SOURCE: Bryan and Shimkin (1943).

These curves were generated by administering three chemicals (individually) to groups of 20 mice and observing the incidence of cancerous tumors for various doses.

Figure 2.1—Classic Example of Dose-Response Curves for Three Chemicals That Can Cause Cancer

explosion in the presence of a human is considered acceptable or without consequences.

The third stage of risk assessment is the exposure assessment. In this stage, the risk assessor must develop a model that indicates the likelihood of individuals coming in contact with the contaminant of concern. For contaminants that have dispersed in soil or water, this evaluation generally takes the form of a transport model showing the movement of the contaminants from their original source location to populated areas. For example, Figure 2.2 shows plumes of contaminants in groundwater at the Massachusetts Military Reservation. As shown, some of these plumes have migrated beyond the boundaries of the installation toward nearby community water supply wells. In contrast, for UXO explosion risk, the exposure assessment must consider the potential for people to migrate to the contaminant. Furthermore, contaminants in soil or water generally disperse in the



SOURCE: Massachusetts Military Reservation Installation Restoration Program,
<http://www.mmr.org/cleanup/index.htm>.

Figure 2.2—Plumes of Groundwater Contaminants at the Massachusetts Military Reservation

form of traceable plumes that affect every molecule of water or soil in contact with the plume. In contrast, for a UXO item, the hazard is isolated to a single point. The exposure assessment must determine the potential for individuals to come in contact with many such points, the precise locations of which are unknown.

It is clear that assessment of risks of UXO explosion will require a different set of tools than has been used traditionally for hazardous waste sites. The nature of risks from UXO explosion is more similar to the risks from the failure of an engineered system—such as explosion of a factory component or an airplane engine—than to risks from chronic exposure to low levels of soil or water contaminants. Like risks from UXO explosion, risks of system failure are binary in nature—either the system fails, or it does not—and the dose-response concept is not relevant. As discussed in Chapters Four and Five, probabilistic risk assessment concepts developed for analyzing system safety could be employed in UXO explosion risk assessment.

Absent formal risk assessment, decisionmaking at UXO sites occurs on an ad hoc basis. Members of a base cleanup team consisting of state regulators, an EPA representative, and Army personnel decide on appropriate actions based on negotiations about what level of clearance must be carried out. When disagreement occurs, negotiations continue until a compromise is reached. Regulatory officials at one location described the appropriate compromise as the “soft pillow” solution: the point at which instinct suggests that the public will be safe and the regulator therefore will not lose sleep. These negotiations occur on a location-by-location basis, and thus the required amount of clearance and the process for deciding upon that amount vary not only from installation to installation but also among locales within a single installation. Risk assessment cannot substitute for negotiations among stakeholders, and in fact stakeholder input is a critical part of the risk assessment process (National Research Council, 1996). However, risk assessment would enable all the stakeholders to be better informed about the nature of the threat and the options for reducing it. Having a standard risk assessment process could lead to more efficient negotiations and a more systematic process for deciding on cleanup options.

CRITERIA FOR AN EFFECTIVE RISK ASSESSMENT METHOD

The Army has recognized the potential usefulness of risk assessment in UXO response. It has tried to develop UXO risk assessment methods but has been unable to garner stakeholder support for these methods. In the next chapter, we evaluate the strengths and limitations of the existing Army methods. The purpose of the evaluation is to determine technical strengths and weaknesses in the Army’s risk assessment toolbox and to identify issues that need to be addressed to increase the credibility of the Army’s analyses in the context of transparent public decisionmaking.

To provide consistency in our evaluations, we developed criteria that an ideal method should satisfy. A number of previous efforts have been made to develop technical criteria for evaluating risk assessments (see Haimes, 1998). The most prominent of these was by Fischhoff et al., initially for Oak Ridge National Laboratory and later published separately in a book entitled *Acceptable Risk* (Fischhoff et al., 1981, pp. 120–128). Other criteria include those developed by the

National Research Council in the report *Understanding Risk: Informing Decisions in a Democratic Society* (National Research Council, 1996, pp. 100–101) and by the Nuclear Waste Technical Review Board, a presidential commission charged with evaluating the technical credibility of the Department of Energy's (DOE) scientific and engineering studies (including its risk assessments) for the proposed Yucca Mountain nuclear waste disposal site (Nuclear Waste Technical Review Board, 2002). We referred to these existing benchmarks as we developed criteria specific to UXO risk assessment. Our criteria were peer reviewed by prominent experts in risk assessment and risk communication as well as by the EPA, and we revised them according to the comments received.²

Table 2.3 shows the criteria and documents the sources from which they were drawn. We grouped the criteria into three categories:

1. **Risk calculation features.** These criteria pertain to the soundness, from a scientific and technical perspective, of the models and risk quantification procedures used by the method.
2. **Implementation features.** These criteria address the adequacy and reliability of the application of the method.
3. **Communication features.** These criteria represent the need to ensure that results can be communicated effectively, so that stakeholders trust the underlying computations and the resulting output.

The complexity of the risk assessment method will of necessity vary substantially depending on whether it is part of a process for programmatic priority setting or detailed site assessment. Nonetheless, the criteria are sufficiently general that they apply to all categories of risk assessment methods. A method that satisfies these criteria would be technically defensible and would provide a means for improving communication with stakeholders.

²Peer reviewers were Dr. B. John Garrick of Garrick Consulting, who was one of the early developers of probabilistic risk assessment; Dr. H. Keith Florig of Carnegie Mellon University, an expert in risk communication; Dr. D. Warner North of Northworks, an expert on decision and risk analysis related to toxic substances in the environment; and James Woolford of EPA, director of the agency's Federal Facilities Restoration and Reuse Office. While comments from these reviewers were considered, the authors accept full responsibility for the criteria.

Table 2.3
Criteria for Evaluating UXO Risk Assessment Methods

Category	Criteria	Sources
Risk calculation features	Is the method comprehensive? That is, does it cover all elements of the risk problem? Is the method logically sound?	Fischhoff et al. (1981); Fischhoff et al. (1981); Nuclear Waste Technical Review Board, 2002
	Is the method consistent with state-of-the-art scientific knowledge?	National Research Council (1996)
	Are the models used for the risk calculations well defined and, ideally, validated by testing against experimental results and observational data?	National Research Council (1996); Fischhoff et al. (1981); Nuclear Waste Technical Review Board (2002)
	Are assumptions explained clearly, used consistently, and checked for reasonableness? Are judgements embodied in the assumptions made clear? Are unnecessary assumptions removed from the final analysis?	National Research Council (1996); Fischhoff et al. (1981); Nuclear Waste Technical Review Board (2002)
	Are uncertainties described accurately and meaningfully?	National Research Council (1996); Fischhoff et al. (1981); Nuclear Waste Technical Review Board (2002)
	Is the output of the tool reproducible? That is, if two assessors have identical information, will they produce the same output with the tool?	National Research Council (1996); Fischhoff et al. (1981)
	Are data sources and calculations presented so they can be checked by others?	National Research Council (1996); Fischhoff et al. (1981)
	Was the method scientifically peer reviewed?	RAND

Table 2.3—continued

Category	Criteria	Sources
Implementation features	Is the method practical for its intended use (priority setting or baseline risk assessment)? That is, are required input data reasonably available or obtainable? Is the level of effort necessary for learning and applying the method reasonable for the intended use?	Fischhoff et al. (1981); Nuclear Waste Technical Review Board (2002)
	Are technical instructions for implementing the method clear?	RAND
	Is the method adaptable to different site conditions and needs? For example, are there provisions for properly addressing risks unique to a particular site?	RAND
	Is the method free of “loopholes” that could be exploited to manipulate results?	RAND
Communication features	Does the method provide specific points for stakeholder input about scenarios considered and assumptions used?	National Research Council (1996); Fischhoff et al. (1981)
	Is the method sufficiently transparent to allow stakeholders to understand the process? Can the results be communicated and understood?	National Research Council (1996); Fischhoff et al. (1981)
	Is the level of uncertainty in results clearly communicated (for example, as a probability distribution)?	National Research Council (1996); Fischhoff et al. (1981); Nuclear Waste Technical Review Board (2002)
	Is the method conducive to educating stakeholders? Does it help stakeholders better understand the problem? Would application of the method help to reduce opportunities for obstructionism?	Fischhoff et al. (1981)

CONCLUSIONS

Although the use of risk assessment at UXO sites has been controversial, a technically sound risk assessment method developed with substantial stakeholder input could improve the quality of decisions at UXO sites and the response selection process. The remainder of this report evaluates the effectiveness (according to the criteria in Table 2.3) of existing UXO risk assessment methods and possible ways of improving future assessments.

It is important to keep in mind that even the best-designed risk assessment cannot eliminate controversy over a decision. Risk assessment can help to structure discussions about which UXO response option—from among all of the imperfect possible choices—is most acceptable to those concerned, but disagreements will inevitably remain over what amount of risk is acceptable from an ethical perspective. In its 1983 review of risk assessment in the federal government, the National Research Council concluded that “Because risk assessment is only one element in the formulation of regulatory actions, even considerable improvements in risk assessment cannot be expected to eliminate controversy over those actions.” Nonetheless, as Fischhoff et al. (1981) point out, “rejecting all [risk assessment] approaches means accepting the marketplace or raw politics, with all their attendant dangers, as the decisionmaking process.” Currently, decisions about level of clearance at UXO sites are made on an ad hoc basis, with a high degree of variability among sites. Employing risk assessment may help bring greater consistency, transparency, and rationality to decisionmaking at UXO sites.

Chapter Three

**EVALUATION OF EXISTING METHODS
FOR UXO RISK ASSESSMENT**

Over the past decade, the Army has developed a number of tools for assessing risks at UXO and other munitions sites. These tools are as follows:

- Interim Range Rule Risk Methodology (IR3M),
- Ordnance and Explosives Cost-Effectiveness Risk Tool (OECert),
- Risk Assessment Code (RAC),
- Ordnance and Explosives Risk Impact Analysis (OERIA), and
- Natural and Cultural Resources Bank (NCRB).

This chapter evaluates these tools. The purpose of the evaluation is to identify strengths and limitations of the existing tool set to help the Army set priorities for future investments in developing credible UXO risk assessment methods.

OVERVIEW OF EXISTING METHODS

Each of the risk assessment methods reviewed in this chapter was designed for a different purpose. Consequently, the methods have different features. Evaluating them is not possible without first understanding the spectrum of uses for which they were intended and the variations in their design. We developed a matrix to categorize the methods according to intended use, types of risks addressed, amounts of data required, and form of output produced. Table 3.1 categorizes the tools according to these features.

Table 3.1
Features of UXO Risk Assessment Methods

	IR3M	OECert	RAC	OERIA	NCRB
Intended uses					
Prioritization		x	x	x	x
Site-specific risk assessment	x	x			
Source(s) of risk addressed					
Explosions	x	x	x	x	NA
Other constituents	x				NA
Data required					
Archival	x	x	x	x	x
Site sampling	x	x		x	
Form of output					
Quantitative		x			
Ordered categorical	x		x	x	x

NA = not applicable.

Intended Uses

As shown in Table 3.1, the methods were designed for two risk assessment applications: prioritization and site-specific assessment.

Prioritization. Two of the methods, RAC and OECert, were designed for program managers to use in allocating their budgets for response to contaminated sites. The output is intended to provide a rough indication of the relative risk of the UXO site, compared with other sites, so that funds and other resources can be directed at the riskiest sites first. A further motivation for prioritization methods is to counter criticism that past UXO response actions have been politically driven, with funding going first to facilities with the most powerful representation in Congress or where citizen concerns were most forcefully voiced.

Site-specific risk assessment. Other methods are designed for detailed, site-specific risk assessments. These methods are intended to provide field managers and concerned stakeholders with information about the nature and significance of the risks at the site in question. They are also intended to help compare the effectiveness of dif-

ferent response options in reducing risks. Ideally, such methods would quantify risk in terms of the number of expected adverse health effects or fatalities per year due to UXO. However, the methods discussed in this chapter do not provide this degree of specificity.

Risk Sources, Data Requirements, and Output

The methods also address different types of UXO-related risks. Most address only the risk from explosions. One method (IR3M) includes both explosion risks and munitions constituents risks.

As shown on Table 3.1, the methods have different data requirements. Some require only archival data, rendering site sampling unnecessary or optional. Archival data are adequate only when the uncertainty in the data will not affect the outcome. In such a case, more data and evidence are unneeded. Other methods require extensive site sampling.

The bottom row of the table shows the different types of output from the methods. Quantitative outputs provide an indication of the degree of risk on a continuous scale. An example is the number of deaths or injuries expected per year due to UXO explosions (although none of the methods we reviewed provides this degree of quantification). In contrast, ordered categorical outputs will group sites into a discrete number of ordered categories, each of which represents a different risk range. For example, sites in category A are less risky than sites in category B, and so on.

CRITERIA FOR AN EFFECTIVE METHOD

Evaluating the available UXO risk assessment tools required us to consider carefully what features are necessary for a tool to be effective for the purposes for which it was designed. To guide our reviews, we developed evaluation criteria in the three categories described in Chapter Two:

- **Risk calculation features.** These criteria pertain to the soundness, from a scientific and technical perspective, of the models and risk quantification procedures underlying the method.

- **Implementation features.** These questions address whether the method can be implemented adequately and reliably.
- **Communication features.** These criteria address the need to ensure that results can be communicated effectively, so that stakeholders trust the underlying computations and the resulting output.

For each category, we applied the list of questions shown in Table 2.1. Not all questions apply to all methods. For instance, methods used only for prioritization will have requirements different from those for site-specific risk assessments. Likewise, a given method may have noteworthy features not addressed by these questions. We used the questions to guide our evaluations but were not rigidly bound by them. Nonetheless, a method that satisfies all of the criteria would have the features necessary to produce scientifically credible output that can be effectively communicated to stakeholders.

The following sections provide summary information about the five risk assessment methods and the key observations from our evaluations of the methods according to the criteria in Table 2.1.

INTERIM RANGE RULE RISK METHODOLOGY

Key Features

The primary technical components of IR3M were designed for use in conducting site-specific risk assessments and in selecting responses. The technical methods address risks from explosion and from munitions constituents. The explosion risk assessment can be based almost entirely on archival data. Determination of risks from munitions constituents requires collecting environmental data and conducting a detailed risk assessment according to EPA procedures. Output from the technical tools is ordered categorically: explosion risks are given a value on a scale of A through E, with E indicating the highest risk; munitions constituents are given values on a scale of 1 through 5, with 5 signaling the highest risk.

Description

IR3M was intended to serve as a DoD-wide standard for risk assessment at UXO sites. It was to provide the technical basis for implementing UXO response policies outlined in the Range Rule. When the Range Rule was shelved, so was IR3M, although efforts have been under way to resurrect IR3M and minor variants with different names.

IR3M includes more than technical tools for risk assessment. It was intended to serve as a complete guide for organizing the UXO response process. The technical instructions for IR3M are extensive: 139 pages plus nearly an equal amount of appendixes. Much of the instructions consist of worksheets for the response manager to complete at every step of the process. The worksheets cover a wide range of information types: from listing people involved in decisionmaking to input data for the risk assessment tools.

Several “offspring” of IR3M have been developed for use at specific sites. For example, at the Adak Naval Air Facility, site managers and regulators have developed the “Adak Island Explosive Safety Hazard Assessment” (ESHA) method. ESHA uses the same approach and framework as the IR3M. The primary difference is that instead of using actual UXO density data, the ESHA method divides sites into three qualitative density categories: (1) UXO not found, (2) UXO strongly indicated to be present, and (3) UXO found (Malcolm Pirnie, 2001).

Risk Calculation Method

The technical portion of IR3M consists of the following tools:

- the “Explosives Safety Risk Tool,”
- the “Other Constituent Risk Tool for Human Health,”
- the “Ecological Risk Tool,” and
- the “Comparative Analysis Tool.”

The remainder of this evaluation focuses on the strengths and limitations of the first three tools. The other components of IR3M, including the extensive set of worksheets and the Comparative Analysis

Tool, are intended to guide the response process and thus do not support risk assessments per se. Therefore, we do not review them further.

Explosives Safety Risk Tool. The Explosives Safety Risk Tool is intended to provide a relative indication of the magnitude of risk from detonation of UXO items. The output is a “risk ranking” on a scale of A to E. The A to E value is determined from a set of decision rules that consider three input factors:

1. **Accessibility (A).** Considers the UXO depth, the potential for it to rise to the surface if buried, and the expected depth of civilian activities such as excavation at the site.
2. **Overall hazard (O).** Considers the type of UXO, whether the fuze is armed, and the amount of energetic material it contains.
3. **Exposure (E).** Considers the UXO density, frequency of entry to the site, intensity of human activity, and weight of the UXO.

Figures 3.1 to 3.3 show the rules for determining A, O, and E. Figure 3.4 shows the rules for assigning a risk value to the site based on the values for A, O, and E.

Sites in category A are deemed the least risky, while those in category E are deemed the most risky.

The Explosives Safety Risk Tool is based on consideration of the main factors that determine explosion risks from UXO: the severity of the hazard posed by the UXO (i.e., the likelihood that it will detonate and the consequences of the detonation) and the likelihood of exposure to UXO. However, these factors are poorly quantified, and they are combined according to what appears to be an arbitrary set of decision rules. The output is not based on reproducible modeling of exposure and hazard.

Further, the Explosives Safety Risk Tool is built on a large set of assumptions about the relative degree of risk posed by different site scenarios. For example, the designers assigned different input values for UXO that are not portable, UXO that can be moved by a vehicle, UXO that can be moved by two adults, UXO that can be moved by a single adult, and UXO that a child can pick up. The basis for the dif-

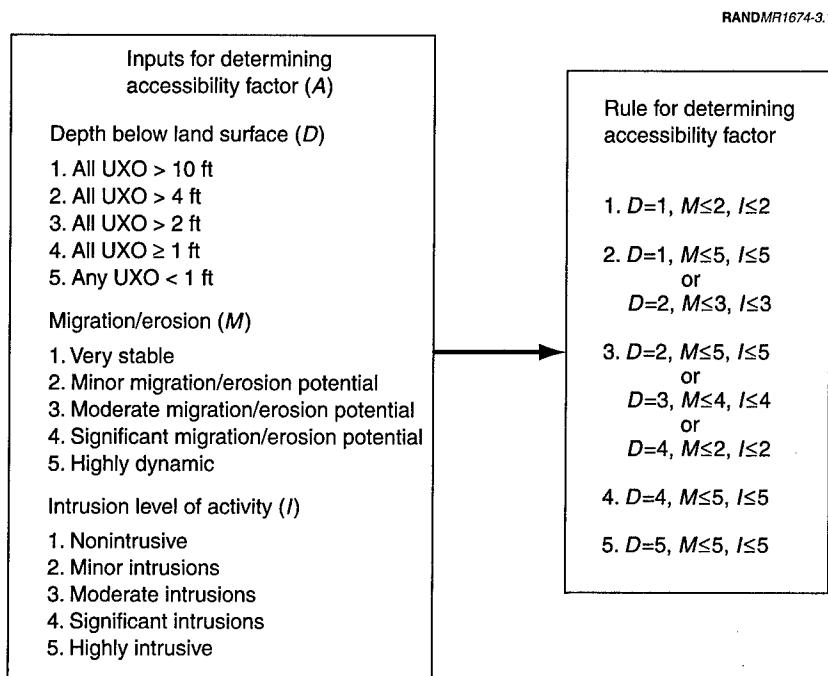


Figure 3.1—Rules for Determining the Accessibility Factor in the IR3M Explosives Safety Risk Tool

ferent assigned input values is not explained. One might wonder how significant the difference in risk is for UXO that can be carried by two adults in comparison to UXO that can be lifted by a single person. The basis for the assumptions is not explained in the IR3M documentation.

The tool does not address uncertainties, either qualitatively or quantitatively. This is important because the output is highly sensitive to some of the inputs, and small variations could have significant effects on the final score. The extensive use of archival data and professional judgment means that the uncertainty in the IR3M output is potentially very large. Use of archival data is acceptable from a technical perspective when uncertainties in a particular input do not

RANDMR1674-3.2

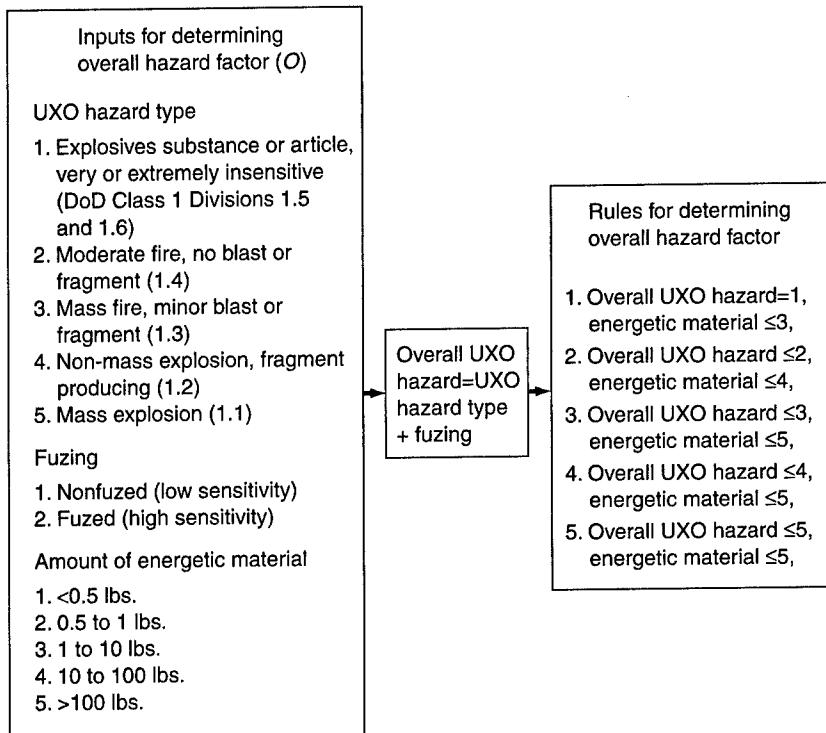


Figure 3.2—Rules for Determining the Overall Hazard Factor in the IR3M Explosives Safety Risk Tool

affect the decision to be made. For example, one data element may have several orders of magnitude of uncertainty, but this uncertainty may have no effect on the final output. However, this is not the case with IR3M. In fact, the output is highly sensitive to inputs that could have large uncertainty ranges due to the use of archival data and professional judgment. The lack of uncertainty quantification and the potentially high sensitivity of the IR3M output to changes in input data mean that the method is not useful for comparing the effectiveness of different response alternatives in reducing risk.

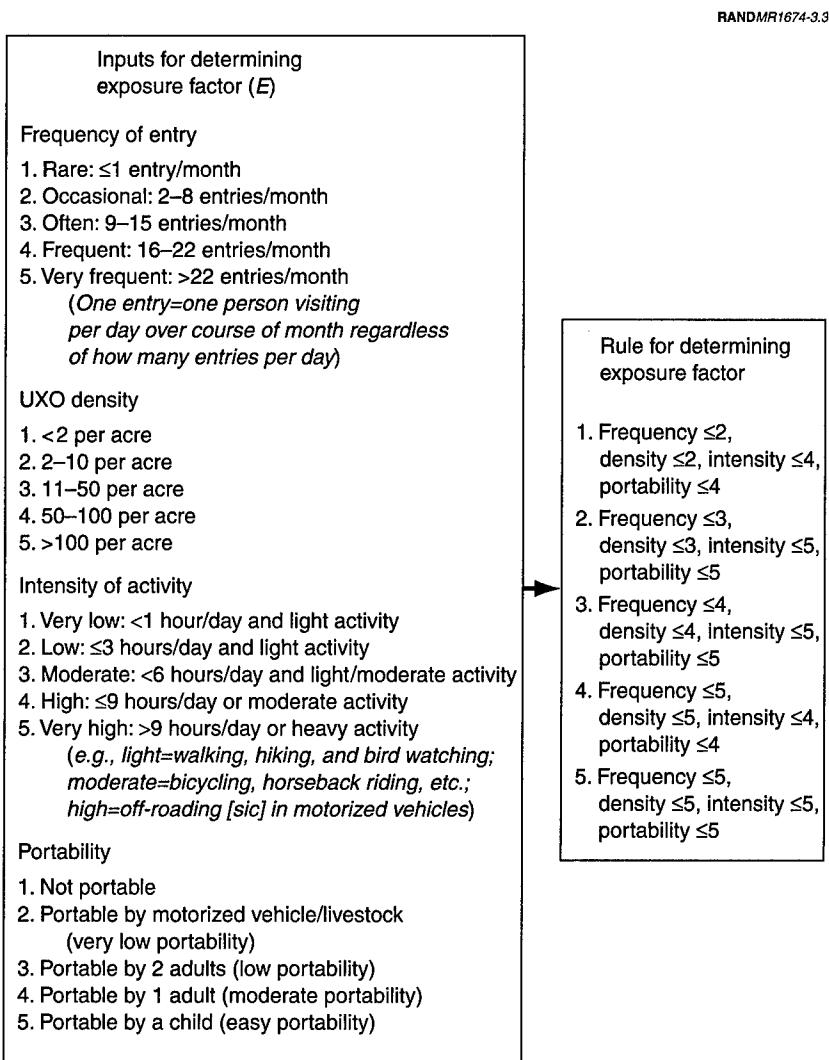


Figure 3.3—Rules for Determining the Exposure Factor in the IR3M Explosives Safety Risk Tool

RANDMR1674-3.4

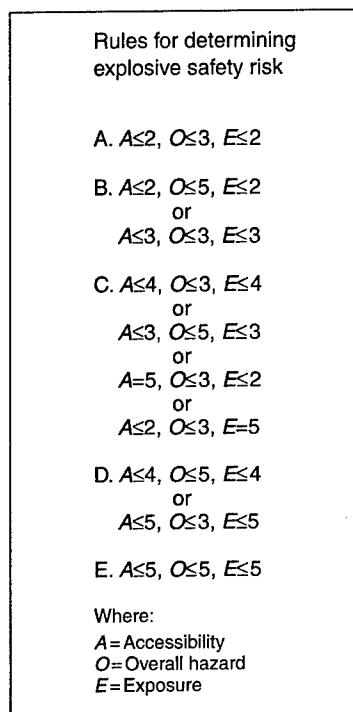


Figure 3.4—IR3M Rules for Determining the Explosives Safety Risk of a UXO Site, on a Scale of A to E

As an example, consider a hypothetical site with the following characteristics:

- Climatic conditions are such that buried UXO is unlikely to surface.
- The site will become a park, and therefore significant digging is not expected.
- The type of UXO present is such that the IR3M “overall hazard” (O) factor equals 5.
- The UXO density, frequency of entry into the park, and UXO weight are such that the IR3M “exposure” (E) factor equals 2.

Now, consider two very similar response options for the site:

- Option 1: excavate all UXO that is less than 0.99 feet deep.
- Option 2: excavate all UXO that is less than 1.01 feet deep.

For all practical purposes (given errors in measuring the excavation depth), these two alternatives are the same. However, IR3M output would indicate that option 2 is vastly preferable to option 1. Following the IR3M rules shown in Figure 3.1, the IR3M “accessibility” factor (A) for option 2 is 3, while for option 1 it is 5. Then, following the rules in Figure 3.4, the IR3M explosion risk value for option 1 is E , while for option 3 it is C . Thus, based on IR3M results, it would appear that excavating an additional 0.02 feet (a quarter of an inch) of soil cuts the explosion risk nearly in half. Intuitively, this output makes no sense. It is not even possible under ordinary operational scenarios to measure the excavation depth with such a fine degree of precision.

An additional flaw of the output of the Explosives Safety Risk Tool is that it is not correlated to known levels of risk. That is, the risk manager has no way to determine whether an option with an IR3M value of E is 10, 100, 1,000, 10,000, or more times as risky as an option with a value of A . The risk tools for other constituents suffer from the same flaw. Although this type of approach might be useful for prioritization, it does not provide enough information to be meaningful for site-specific risk assessment.

Other Constituent Risk Tool for Human Health. The Other Constituent Risk Tool for Human Health uses a process similar to that of the Explosives Safety Risk Tool, except that the output is on a scale of 1 to 5 rather than A to E , with 1 indicating the lower risk. Figure 3.5 shows the inputs for the human health risk tool and the rules for determining the risk value. IR3M states that the cancer risk and hazard index values are to come from detailed site investigations carried out according to EPA’s Risk Assessment Guidance for Superfund (RAGS) process. Further, it indicates that only factors for which information is available need to be considered. For example, if no information is available that would allow determination of blood lead levels, then this factor does not need to be considered in the evaluation. IR3M does not provide guidelines other than those shown in Figure 3.5 for determining which “other factors” are important and whether these

RANDMR1674-3.5

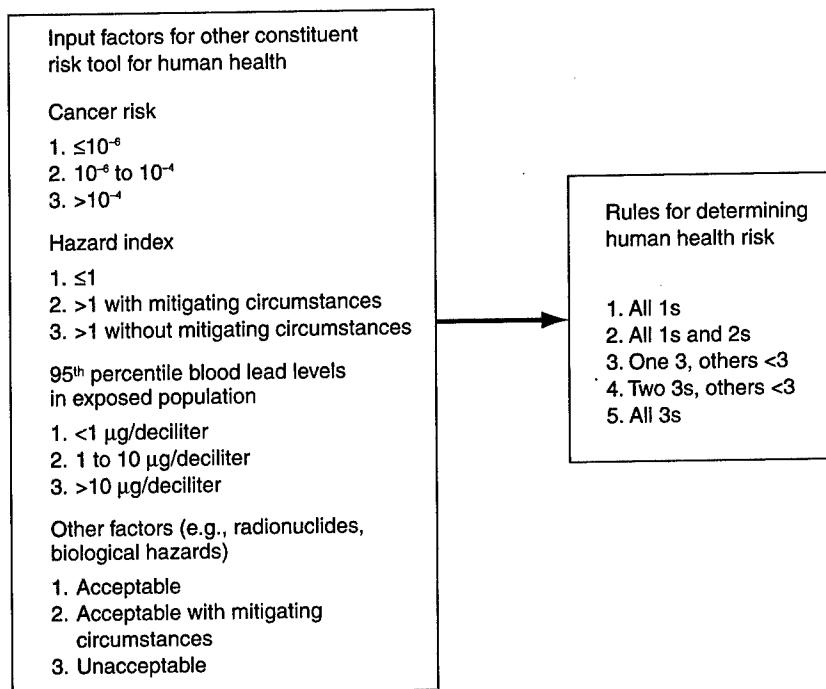


Figure 3.5—Input Factors and Rules for Determining the IR3M Value for Human Health Risk

factors are “acceptable.” Also, IR3M does not indicate how to decide whether “mitigating circumstances” are present when determining an appropriate value for the hazard index input factor.

This tool raises another major flaw in IR3M: the output may not be reproducible. Two evaluators given the same set of site data could compute different risk levels, depending on their interpretations of the input factors. As an example, consider a UXO site with the following characteristics:

- TNT has migrated from degraded UXO into groundwater.
- TNT has been found in a nearby community drinking water well at a concentration of 0.32 ppm.

- The Army has supplied bottled water to the community.
- Monitoring has determined that the well water contains no lead or other munitions chemicals.

For TNT, the primary health concerns are noncancer effects, including anemia, aplastic anemia, hepatitis, and urine discoloration. Although EPA has categorized TNT as a possible carcinogen, data are not available to allow computation of cancer risk levels. Thus, in this case, the relevant consideration is noncancer risk.

Noncancer risk assessment is based on the concept of "acceptable daily intake," also called the "reference dose." The reference dose represents the amount of a contaminant that all humans, including sensitive populations (such as children) may be exposed to every day without appreciable risk of adverse health effects. The risk of a contaminant with noncancer health effects is then quantified using a "hazard index," which is the ratio of the amount of a substance to which people are exposed to the reference dose. An index greater than 1 signals the potential for harm and is often a remedial action trigger.

For our example, if we assume the average individual weighs 70 kg and consumes 2 liters of water per day containing 0.32 ppm of TNT, then the individual's daily dose of TNT is 0.01 mg/kg/day. The reference dose for TNT (from the Agency for Toxic Substances and Disease Registry) is 0.0005 mg/kg/day. Thus, the hazard index is

$$\text{hazard index} = 0.01 / 0.0005 = 20.$$

The hazard index is much greater than 1, signaling a high risk if individuals were to consume the TNT-contaminated water.

We could use this information to apply the IR3M human health risk tool. The cancer risk is not relevant; there is no risk of lead exposure; and there are no other factors to consider. Thus, according to the IR3M rules shown in Figure 3.5, the only necessary input for the human health risk tool is the hazard index. The hazard index is much greater than 1, so the hazard index factor for the risk tool would be either 2 or 3, according to Figure 3.5. One evaluator might decide that the provision of bottled water constitutes "mitigating circumstances" and choose a value of 2. However, another might de-

cide that the bottled water supply is temporary and therefore choose a value of 3. Based on these inputs, the final output of the human health risk tool could be either 2 or 3, depending on the judgment made by the evaluator.

A similar example calls into question the logical soundness of the IR3M Other Constituent Risk Tool for Human Health. For this example, assume that RDX, rather than TNT, has migrated into the water supply well at a concentration of 14 ppm,¹ with all other characteristics of the hypothetical site the same as before. In this case, the relevant concern is cancer risk: RDX is a carcinogen, and EPA has set an acceptable dose of 0.005 mg/kg/day, corresponding to a cancer risk level of 10^{-4} . For the hypothetical example, the 14 ppm concentration is equivalent to a daily dose of 0.4 mg/kg/day—80 times higher than the 10^{-4} cancer risk level. Thus, for the hypothetical site, one would select an IR3M cancer risk value of 3, according to Figure 3.5. The human health risk tool output would also be 3 (because, as in the first example, no other factors are relevant). This value of 3 is intended to represent a moderate risk on the 1–5 scale, yet the site warrants a higher risk level because the RDX concentration is so high. Thus there is a flaw in the underlying logic of the tool: a risk level that appears high based on input information produces an output that indicates only a moderate risk rating, according to IR3M.

Ecological Risk Tool. The Ecological Risk Tool is analogous to the Other Constituent Risk Tool for Human Health. Input factors are to be derived in part from a detailed assessment conducted in accordance with EPA's Ecological Risk Assessment Guidance for Superfund (ERAGS). Figure 3.6 shows the input factors for this tool and the rules for choosing a risk value. IR3M does not explain how to determine values for the input factors. For example, it does not provide quantitative definitions for distinguishing “small area” from “medium area.”

¹The Agency for Toxic Substances and Disease Registry reports that RDX concentrations as high as 14 ppm have been found in groundwater.

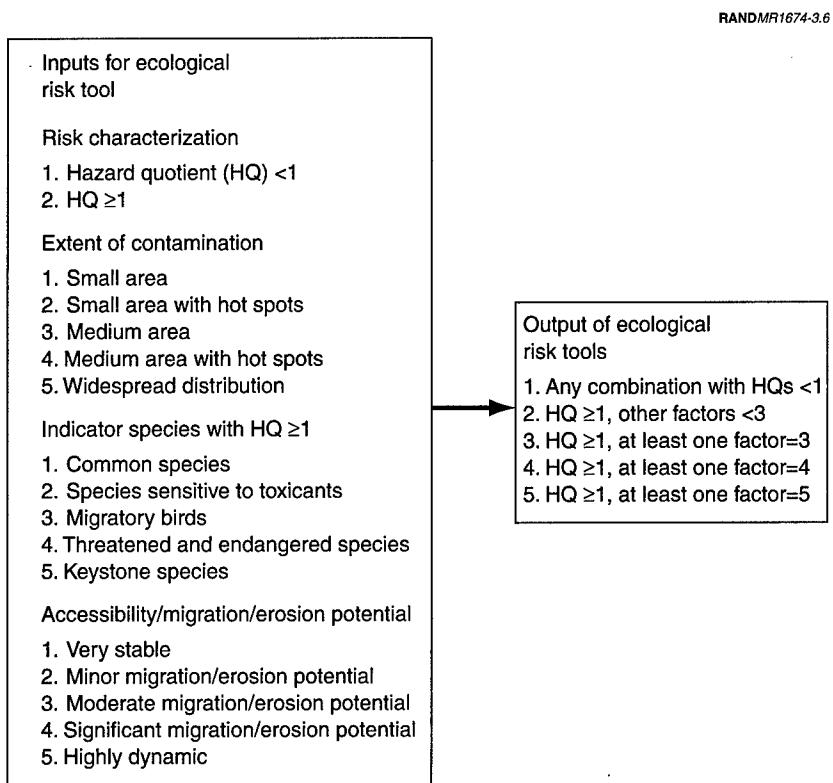


Figure 3.6—Input Factors and Associated Output for IR3M Ecological Risk Tool

Implementation Features

The basis for the IR3M tools is simple (overly so, in our estimation), so in theory it should be easy to implement. However, the instruction manual is not easy to follow. Further, the instructions are not clear enough to ensure that the tool is applied uniformly.

Some of the worksheets are clear and easy to follow, but many are extremely confusing. For example, one worksheet instructs the user to answer the question, “How will decisions be made?” The intent is for the user to specify data objectives for the site evaluation, but no

instructions are provided for developing these objectives other than the following statement: “The DQOs [data quality objectives] should be focused on providing the necessary information to make the required decisions at this point in the process.” Many worksheets contain unique terminology that is not defined in the manual. Examples include “qualitative tolerable error limits” and “judgmental sampling.” An even greater flaw is the lack of specific instructions for choosing input values for the risk tools. For example, as Figure 3.1 shows, in order to compute the accessibility factor for the Explosives Safety Risk Tool, the evaluator must determine whether the UXO is “very stable,” a condition for which the assigned value is 1; subject to “minor migration/erosion potential,” for which the assigned value is 2; subject to “moderate migration/erosion potential,” with an assigned value of 3; or likely to undergo “significant migration/erosion,” with a value of 4. No guidance is given on determining the difference between “minor,” “moderate,” and “significant.” The manual contains numerous typographical errors, which further increase the confusion.

We also have concerns about the level of data required. In theory, the Explosives Safety Risk Tool could be used without gathering site data. The use of archival data in most cases will not be adequate for conducting site-specific risk assessments. In contrast, detailed sampling is required to determine the inputs for the human health and ecological risk tools, which are based on the results of EPA’s RAGS and ERAGS assessment tools. Thus, data requirements for the different tools are highly uneven.

Communication Features

The IR3M output cannot be easily communicated to stakeholders, for several reasons. First, the technical instructions are confusing. Second, stakeholders could easily object to the selection of input values because so many of them are based on subjective judgments and because little or no guidance is provided to standardize judgments. Third, the output is of dimensionless form that does not have a clear correlation to risk and therefore cannot assure stakeholders that their concerns for public safety are being addressed. Finally, stakeholders raised a number of concerns in initial reviews of IR3M and exercises designed to test it that were never addressed. For example,

a team consisting of two DoD representatives, two EPA representatives, and one citizen group representative that was asked to evaluate a hypothetical site using IR3M commented, "Intuitively, the [IR3M] scores did not make sense . . . If action removes all known UXO, the tool does not reflect a change in protectiveness and therefore there is a flaw in the tool" (Department of Defense, 2000).

Conclusions

In conclusion, IR3M has serious limitations that should preclude its use. The method's logic is not sound. It can produce output that masks known risks or inflates risk values. The assumptions used as the basis for many of the calculations are not explained or are not reasonable. The method does not provide any means to account for uncertainty. The output is not always reproducible: two assessors could produce very different risk values using the same IR3M tools. Technical instructions for the method are unclear in many instances. There are many loopholes under which assessors could inflate or deflate risk values by choosing certain assumptions. Data requirements for some portions of IR3M do not reflect the complexity of the underlying problem. And finally, the steps in the process would be difficult for stakeholders to follow, and the output is not useful for communicating risk levels. A previous review of IR3M by the Army Science Board reached similar conclusions (Army Science Board, 1998).

ORDNANCE AND EXPLOSIVES COST-EFFECTIVENESS RISK TOOL

Key Features

The Army Corps of Engineers and its contractor, Quantitech, developed OECert to prioritize UXO sites or sectors within sites using risk alone, cost alone, or by a cost-effectiveness ratio of risk reduction per unit cost. In practice, however, OECert was used as a site-specific risk assessment method, until objections to the method from regulators caused the Army to stop using OECert and create an alternative (IR3M). OECert uses both archival and site-specific data to characterize UXO sites and estimate exposures. OECert quantifies risk by combining these exposure estimates, the population in the

surrounding area, and the density of UXO in the area, assuming a standard hazard characterization of UXO. It does not address risks from munitions constituents.

Description

The measure of risk in OECert is defined as

$$\text{Risk} = (\text{Number of Expected Exposures to OE}) \times (\text{OE Hazard Factor}),$$

where OE is ordnance and explosives. The likelihood that this product accurately reflects the risk at any given site is limited by the credibility of the OE hazard factor, the ability to determine accurately the number of expected exposures (which combines estimations of the number of persons performing specific activities and a measure of the UXO density), and other assumptions. In some cases, the results of OECert component models (e.g., derivation of the OE hazard factor) are not readily interpretable outside of OECert, and this fact is not transparent in the resulting estimation of risk. Such a design approach might be appropriate if OECert were used as a prioritization tool as intended, but it is not adequate for site-specific risk assessment.

An underlying technical requirement of OECert is the ability to divide a UXO site into sectors, each with a spatially homogeneous distribution of OE. OECert classifies a UXO area into one of three sector types for analysis: “dispersed” (e.g., impact areas); “localized” (e.g., burial pits and trenches); or “water” (i.e., dispersed or localized sectors under water, also referred to in OECert as water or shore locations). Homogeneity within a sector is based on factors such as vegetation, terrain slope, soil type, future land use, and UXO density, and it is generally determined using statistical estimation methods.

OECert’s risk calculation also depends on an estimation of exposure. Exposure is based on the consideration of 19 public activities that may occur at a UXO site; the selection of these activities is based on projections of future land use. Table 3.2 shows the activities. The expected number of people performing an activity annually is estimated by multiplying the population in the area by the percent of that population expected to perform the specific activity. The intent of this process is to estimate the area of a UXO site that would be tra-

Table 3.2
Activities Resulting in Potential UXO Exposure
(as Defined in OECert)

Occupational Activities	Recreational Activities
Construction	Children playing
Surveying	Walking shortcuts
Archeological digging	Picnicking
Crop farming	Camping
Ranching	Hunting
	Freshwater fishing
	Offroad driving
	Mountain biking
	Hiking
	Swimming
	Horseback riding
	Motorbiking
	Metal detecting
	Jogging

versed by a population performing the activity. OECert contains flow charts to guide this process.

OECert combines activity information with UXO density information to determine the number of expected exposures. Then, as shown in the equation above, this exposure estimate is multiplied by the OE hazard factor (the basis for which is explained later in this section) to determine risk.

Risk Calculation Method

OECert's risk calculation method varies somewhat depending on the type of UXO site or sector (impact area, disposal area, or submerged area).

Dispersed sectors. OECert assumes that dispersed sectors (e.g., impact areas) have "UXO randomly distributed over a relatively large geographical area." These sectors are defined as the largest geographic area that has homogeneous terrain and UXO density. UXO

density and environmental site data are used to determine homogeneity of the sectors; because OECert was not intended to assess risks from munitions constituents, these data are not used to determine the fate and transport of such constituents, nor are they used to inform a public health or ecological risk assessments. Table 3.3 details which types of environmental data OECert requires and for what purpose.

Following a division of the site into dispersed sectors, OECert estimates the number of exposures an individual will have to UXO in a specific sector during a specific activity (number per activity and number per year). OECert then calculates risk for each sector by applying a UXO hazard factor and adding the sector risks to provide a risk estimate for the entire area.

Localized sectors. OECert's localized sector method (for burial pits and trenches) is very similar to the dispersed-sector method. It appears that the methodologies were divided primarily to allow for differing exposure calculations. OECert divides localized sectors into three categories, each with different exposure and risk estimation factors (see Table 3.4).

Table 3.3

Environmental Data Used by OECert to Determine Homogeneity of UXO Sites or Sectors

Type of Data	Rationale for Consideration
Slope	Clearance is more difficult on slopes, and slopes either may act as obstacles, discouraging public activity, or may reduce the total amount of area traversed per activity time
Vegetation	Lower vegetation density is associated with lower response cost but higher probability of public activity (dense vegetation is an obstacle)
Soil type	Soil type influences UXO penetration depth
Plants and animals	Dangerous animals (e.g., snakes) and plants (poison ivy) will slow response and increase cost; presence of endangered species may stop work altogether

Table 3.4
Categories of Localized Sectors in OECert

Category	Exposure/Risk Estimation Factors
Localized excavation (buried ordnance)	Area of OE contamination Area traversed by personnel performing ground-intrusive activities Number of individuals performing specific activities annually
Localized surface (stockpiles)	Area of OE contamination Line of sight to the contaminated area Area of site Number of individuals performing specific activities annually
Localized building (storage buildings)	Number of buildings Population of the state in which the site is located Public exposure (estimated using the burglary rate in the state)

Despite many similarities, there are some distinctions between the dispersed and localized sector methodologies. In estimating risk, the UXO hazard factors for localized sectors differ from those for dispersed sectors. For example, UXO containing white phosphorus is assigned a hazard factor of 13 for dispersed sectors but a hazard factor of 3 for localized sectors. Also, eight activities, instead of nineteen, are considered for exposure to localized excavation sectors: children playing, offroad driving, picnicking, camping, construction, crop farming, archeological digging, and metal detecting. Breaking and entering (based on local crime records) is the only exposure route considered for UXO in localized building sectors.

Water sectors. Submerged UXO is considered to be either dispersed or localized. Again, the method is generally similar to those for dispersed and localized sectors, although there are some differences. Environmental data collected are the same as for the other methods except that water depth and strength of current (none, moderate, or strong) are included. Different activities, shown in Table 3.5, are used for estimating exposure in water sectors.

Table 3.5
**Activities Considered by OECert when Estimating Exposure
 to UXO in Water Sectors**

Shore	Both	Water
Picnicking/sunbathing	Children playing	Snorkeling
Dune buggy driving	Freshwater fishing	Skiing
Hunting	Saltwater fishing	Diving
Hiking		Boating/sailing
Moped riding		Surfing
Construction		Sailboarding
Archeological digging		Swimming
Metal detecting		

Technical Models and Assumptions

OECert relies on several embedded technical models and underlying assumptions. The technical models include statistical methods to determine homogeneous sectors, a “density analogy process” to predict UXO density based on prior sampling of UXO sites, an “analytic hierarchy process” to derive UXO hazard factors, and multiple exposure models. Each of these requires explanation to understand the strengths and limitations of OECert.

Reliance on homogeneous sectors. The use of homogeneous sectors is intended to ameliorate significant uncertainty associated with UXO density data. As Chapter One explained, UXO detection methods cannot find all buried ordnance and have trouble distinguishing between UXO and anomalies such as background interference due to natural (e.g., magnetic rock) or introduced objects. The high false alarm rates of metal detectors make a complete accounting of UXO at any site difficult or impossible. In addition, many UXO sites are vast in area, and deploying sampling equipment over thousands of acres may be impractical, from the Army’s perspective. For this reason, the Army developed techniques for OECert to estimate UXO density and derive homogeneous sectors.

Density estimates. After OECert was developed, the Army Corps of Engineers recognized the need for a statistical UXO density estimation tool and created a computer program that combines statistical analysis and random UXO sampling to meet this need. The original density estimation tool was called “SiteStats/GridStats.” A newer tool, the “UXO Calculator,” has replaced SiteStats/GridStats. The main difference between these tools is that UXO Calculator provides for use of a digital geophysical mapping device. UXO Calculator also allows the assessor to input a certain UXO density based on future land use and then calculate the number of samples required to determine if that density has been exceeded.

To conduct statistical sampling for UXO Calculator, a grid is located within a (presumed) homogeneous sector (typically 50 × 50, 100 × 100, or 100 × 200 feet). The grid is cleared of vegetation and scanned using a detection device selected for the particular site. Anomalies² are marked. If the detector finds fewer than 20 anomalies within a grid, then the survey crew excavates all anomalies. When the detector finds more than 20 anomalies, then an analyst selects 25 to 33 percent of them for excavation based on a combination of a statistical sequential probability ratio test and ad hoc stopping rules. Once the anomalies are identified, results are fed into a software program. The software then uses principles of random sampling to determine which anomalies to excavate next, which grids to sample next, and so forth. The software determines when an adequate portion of the site has been sampled and the investigation is complete. Finally, based on the investigation of a sufficient number of grids within a sector, the density of UXO is extrapolated to the entire sector (EPA, 2001).

The use of statistical methods to estimate UXO distribution is controversial in the regulatory community. The basis of this controversy is the treatment of UXO distributions as homogeneous, which must be assumed to support the results of random sampling. Statistical sampling relies on the assumption that the probability of detecting UXO in one location is the same as in another location. However, this assumption has not been validated convincingly. If the distribution of UXO is not truly homogeneous, then statistical sampling

²Here, an anomaly refers to the detection of any object—whether UXO or natural or man-made clutter.

methods could underestimate or overestimate UXO occurrence. It is unlikely that UXO is uniformly distributed over an area of concern rather than concentrated in former impact areas and target locations. The Army Corps of Engineers has attempted to distinguish between such uniform distribution and “spatial homogeneity” of UXO, but this distinction has not been explained to the satisfaction of regulators.

Regulators have expressed four additional concerns about statistical sampling practices for UXO: (1) the inability of site personnel to demonstrate that the assumptions of statistical sampling have been met, (2) the extrapolation of statistical sampling results to a larger area without confirmation or verification, (3) the use of the density estimates in risk algorithms to make management decisions about the acceptable future use of the area, and (4) the use of statistical sampling alone to make site-based decisions. In general, regulators suggest that statistical sampling is best used as a screening tool, with additional samples gathered as site investigation proceeds (EPA, 2001).³

Density analogy estimates. OECert includes a “density analogy process” for use in lieu of statistical sampling. This process is based on “best engineering judgment” density values and comparison to selected other UXO sites. The OECert developers conducted detailed UXO density evaluations for six sites (although only four of the six evaluations were completed in time to be included at the time the OECert documentation was published). When applying OECert to a new site, the assessor has the option of using a “density analogy process” to provide estimates of the UXO density and the percentage of UXO on the surface based on the data collected at the four pilot sites.

For a new sector requiring density estimates, the density analogy procedure is as follows:⁴ The sector is subjected to three binary classifications resulting in eight distinct analogy classes, which are a function only of what kinds of ordnance-releasing activities took place on the site (e.g., training area versus proving ground, bombing versus firing) and for how long (number of years). The classification

³EPA (2001) provides a more thorough critique on the use of statistical sampling methods, beginning on page 7-29.

⁴The process and associated formulas are included in OECert Appendix F.

does not account for any other physical characteristics of the site such as terrain and weather, nor does it account for the specific types of munitions used at the site. After deciding to which of the eight classes the new sector belongs, its “analogous densities” are produced by averaging the corresponding quantities for all sites in the database that are in that same class. At present, because so few detailed evaluations have been used to build the database, this average is over only a small number of sites.

The justification for the mechanics of the averaging deserves some discussion. The simplest case is the percentage of UXO on the surface, in which the site analogy is simply the arithmetic mean of these percentages for all relevant sites in the database. The rationale is that this percentage should not depend on the number of years the site was used. The more questionable case is the derivation of the analogous UXO density that is weighted by years, the justification being that density should increase as the number of years the site was used increases. On its face, this makes some sense, but the (albeit) limited data included in the tool may indicate otherwise. For example, the Camp Croft site was used for three years, but its UXO density is approximately eight times higher than the Mission Trails site, which was used for eight years. At the very least, even if the assumption is true, other factors could overwhelm this effect with the amount of data used to support the model.

Moreover, it is not clear that the formula used to estimate the UXO density is appropriate. It is sensitive to outliers, especially due to the small size of the database. For example, if one augments the database with “clean” sites that have never been used—that is, they have zero UXO density—then the resulting analogous UXO density for the site being assessed can depend on the number of such clean (or nearly clean) sites in the database.

The current database is too sparse and variable to have much practical value in supporting the density analogy process. It is not clear that the calculation of analogous quantities derived from these data would be any better than an estimate based on expert judgment; in fact, the expert judgment could account for important factors not

currently considered, such as UXO types present and environmental characteristics of the site.⁵

Hazard factor. The calculation of risk in OECert depends not only on exposure estimations but also on the OE hazard factor.⁶ The OE hazard factor is derived from a formal approach known as the “analytical hierarchy process.” In this approach, an expert compares pairs of K different hazards, in each case providing some subjective judgment about which hazard is worse and by how much. OECert uses $K = 11$ for the different classifications of UXO, and for each of the 55 possible pairs of UXO types, experts were asked to rate the relative hazard of type i to type j on an integer scale of 1 to 9, with a value of 1 indicating equal hazard and a value of 9 indicating type i being maximally more hazardous than type j . After some unspecified calculations, these results were used to produce a single number (“weight factor”) for each UXO type. This weight factor presumably conveys the relative hazards of the classes. The discussion of this process in OECert includes the resulting weight factors for two different groups of experts. Despite a few notable discrepancies, the weight factors elicited from the two groups are generally consistent.

To carry out the OE hazard factor calculation, experts were asked to make pairwise comparisons for the 11 classes twice. The first time was to compare detonation sensitivity; and the second time was to compare detonation consequences. For each of the 11 classes, two quantities (sensitivity and consequences) were multiplied together, and the “adjusted hazard factor” was calculated by dividing each product by the collective maximum over the 11 classes. This resulted in a scale of 0 to 100, with larger values indicating larger hazards.

Separating the components of detonation sensitivity and detonation consequence is reasonable; for a few classes of UXO, the two scores may significantly differ. However, the justification for multiplying together the two scores to derive the unadjusted hazard factor is unclear.

⁵An Army Corps of Engineers official told the authors that to date, the density analogy tool has not been used in lieu of site sampling.

⁶OECert Appendix G describes the process for determining the hazard factors.

After multiplying the OE hazard factor by the number of exposed individuals, it seems that the resulting number may be a reasonable indication of relative risk, although the resulting value has no meaning outside of OECert.

Exposure estimates. Most of the main body of the OECert documentation is devoted to calculating estimates of exposure.⁷ This level of detailed modeling is a strength of OECert if it is appropriately applied.

OECert estimates probabilities of exposure to buried UXO based on the spatial area used for particular activities and intrusion depths for those activities (e.g., each picnicker is estimated to intrude up to 1 foot in a 0.25-square-foot area). Individual exposure is calculated as the probability of a person being exposed to UXO during a given activity based on a model of that activity. Public exposure is the sum of all the individual exposures and applies to the population in the area.

It seems clear that the developers of OECert created the exposure models in an attempt to provide conservative estimates of risk (i.e., to err on the side of overestimating exposure). However, these models, like many exposure models, rely on a large number of assumptions. In some cases, the exposure assumptions are based on consultations and data. For example, for crop farming, the number of times a farmer must traverse a field to cultivate a particular crop was calculated based on the experience of agricultural agencies. In many other cases, estimates are based on data such as demographics and the amount of sunshine per year (a factor for recreation). In most cases, however, the assumptions do not appear to have been based on consultations with stakeholders or to have been validated by observation. Regardless of how the assumptions were derived, the ultimate activity value is still an estimate and should have some uncertainty related to it, the magnitude of which might be adjusted based on observation.

OECert does not address the possible variability of public behavior. Even though the intent might have been to overestimate exposure, the credibility of the exposure estimates may be questioned without some explanation of the uncertainty or measurements designed to

⁷This methodology is further explained in OECert Appendix H.

validate the estimates. Further, the lack of explicit treatment of uncertainty weakens the basis for comparisons between sites and any claim of “true” characterization of risk.

Implementation Features

Use of OECert has declined for several reasons. Stakeholders have objected to its statistical tools (SiteStats/GridStats and UXO Calculator) for determining sampling requirements and estimating UXO densities. The technical instructions in OECert are generally, but not always, clear. Similarly, guidance to the user is lacking in some areas. For example, for collecting site data to determine homogeneity of sectors, OECert provides guidance for classifying terrain slope (less than 10°, 11 to 30°, or over 30°), but it lacks similar specific guidance for vegetation, soil type, and land use. Similarly, the models and assumptions that support OECert are difficult to understand because they are often not clearly explained.

OECert relies on extensive data collection, especially of archival data. The data-collection process is guided by software (Excel® and Visual Basic templates). In our review of OECert, we attempted to create a hypothetical site investigation using this software, but the application crashed repeatedly. It is not clear whether this experience reflects actual experiences using OECert in the field, but the amount of information required by OECert suggests that software reliability is an important consideration.

Communication Features

OECert does not include explicit provisions for stakeholder involvement in the evaluation process, nor are its methods readily transparent to stakeholders. OECert exemplifies the tradeoffs between a complex and (potentially) precise tool and the requirement that methods be simple enough to communicate to stakeholders. While a point estimate of risk such as is produced by OECert might seem easy to communicate, it can also convey a false sense of accuracy while masking uncertainty and the actual range of potential risks. At the root of OECert’s difficulty in communication are its embedded and often opaque models and assumptions.

Conclusions

OECert was originally intended to support structured, risk-based decisionmaking to prioritize response efforts among and within UXO sites. However, it was perhaps too complex for prioritization and was used instead for site-specific assessments until use was curtailed due to regulatory concerns. The overall OECert structure suggests that with some modification, it could serve as a starting point for developing a more robust UXO site characterization tool. However, it is unlikely that it will serve as a credible risk assessment tool without further modification, including, for example, improved treatment of uncertainty, improved processes for determining UXO density, consultations with stakeholders in developing exposure models, and validation of exposure models.

RISK ASSESSMENT CODE

Key Features

RAC, developed by the U.S. Army Engineering and Support Center in Huntsville, Alabama, is a prioritization method designed originally to determine the relative risk of UXO explosions at FUDS (Army Engineering and Support Center, 2000). The method provides an ordered categorical ranking of the urgency of UXO explosion risks on a scale of 1 through 5, with 1 indicating the highest-priority sites for action. The ranking is based on a simple algorithm that considers the severity of the UXO hazard and the probability that humans will come in contact with UXO. RAC addresses explosion risk only. It does not address human health and ecological risks from munitions constituents. Instructions state that input data should be obtained from archival sources and limited site sampling: “The risk assessment should be based on the best available information resulting from record searches, reports of Explosive Ordnance Disposal detachments . . . , field observations, interviews, and measurements.”

Description

For this tool, the evaluator develops overall scores for indexes of hazard severity (maximum value 61) and hazard probability (maximum value 30) by adding up scores for a number of subfactors.

The hazard severity value is determined by inventorying all types of ordnance present and assigning values to each type, regardless of the number present. Scores for five categories of ordnance are summed to obtain the total hazard severity score. The five ordnance categories are (1) conventional, (2) pyrotechnics, (3) bulk high explosives, (4) bulk propellants, and (5) chemical and biological. RAC includes worksheets to guide the scoring. As an example, Table 3.6 shows the inputs for the conventional ordnance score. The evaluator is instructed to circle all types of conventional ordnance present and choose the largest single value. Similar scores are developed for pyrotechnics, bulk high explosives, bulk propellants, and chemical and biological warfare materiel.

The hazard probability value is obtained by summing up scores for the area, extent, and accessibility of the UXO hazard. As for the hazard severity value, RAC includes worksheets for scoring these inputs. The hazard severity and hazard probability values are combined on a matrix to determine the priority for action. Table 3.7 shows the matrix. A final score of 1 indicates the highest priority for action and 5 the lowest.

Table 3.6
RAC Hazard Severity Values for Conventional Ordnance

Ordnance Type	Score (points)
Medium/large caliber	10
Bombs, explosive	10
Grenades, hand or rifle, explosive	10
Landmine, explosive	10
Rockets, guided missile, explosive	10
Detonators, blasting caps, fuzes, boosters, bursters	6
Bombs, practice (with spotting charges)	6
Grenades, practice (with spotting charges)	4
Landmines, practice (with spotting charges)	4
Small arms, complete round (.22-.50 caliber)	1
Small arms, expended	0
Practice ordnance (without spotting charges)	0

Table 3.7
RAC Matrix for Determining Priority for Action

Severity Category	Probability Level				
	Frequent (score > 26)	Probable (score 21–26)	Occasional (score 15–20)	Remote (score 8–14)	Improbable (score < 8)
Catastrophic (score > 20)	1	1	2	3	4
Critical (score 10–20)	1	2	3	4	5
Marginal (score 5–9)	2	3	4	4	5
Negligible (score 1–4)	3	4	4	5	5

Risk Calculation Features

RAC uses a simple model, in the sense that its output is based on the following computation:

$$\text{RAC score} = (\text{hazard severity}) \times (\text{hazard probability}).$$

That is, the risk from UXO depends on the likelihood of an individual encountering a UXO item and the magnitude of an explosion if one occurs. Underlying assumptions in the model are explicit in the instructions. RAC explains clearly how scores are determined for hazard severity and hazard probability. For example, one of the input values for the hazard probability factor is the distance to the nearest inhabited location; RAC provides clear instructions about how to convert measured distances to the appropriate input value for RAC. If any structures are less than 1,250 feet from the site, the appropriate input value is 5; if structures are between 1,250 feet and 0.5 mile, the input value is 4; and so on. However, a limitation is that the justifications for these assumptions are not provided. For example, RAC does not explain the basis for determining “safe” distances. A further limitation is the lack of provisions for documenting uncertainty in the input values. An additional and very significant limitation is that the hazard probability score does not account for UXO density. Further, RAC does not consider the amount and combined risks from

the presence of multiple types of UXO: the most hazardous type is chosen as the basis for evaluation, even if the range contains only one such explosive item and the remainder pose no or minimal risks, such as expended small munitions.

Because RAC provides detailed instructions for computing scores, the results are likely to be reproducible. That is, two different evaluators will reach the same conclusion about whether the site should be assigned a high priority for action.

Implementation Features

The instructions for using RAC are clearly written and unambiguous. The tool is easy to use: instructions are concise (8 pages total) but complete. The required input data can be readily obtained from archival information and minimal site investigation, which is appropriate for a tool intended for prioritization only.

Communication Features

RAC does not contain explicit provisions for obtaining stakeholder input, nor does it provide guidance on communicating results. In general, the clarity of RAC's instructions provides a basis for explaining the rationale of the scoring method.

Conclusions

Overall, RAC seems well suited for its intended purpose: prioritization of UXO FUDS sites for response action based on the risk of UXO detonation. The instructions for RAC make it very clear that the tool is intended only for priority setting. However, the tool has been misused for site-specific risk assessments, and, as a result, regulators have criticized it. The tool lacks sufficient detail to serve as the basis for a thorough site evaluation. For example, UXO depth is classified simply as "surface only" or "subsurface," with no differentiation provided for UXO present at different subsurface depths—information that is not necessarily essential for initial priority-setting but that is critical for conducting technically credible site-specific risk assessments. Similarly, as noted above, the RAC score does not con-

sider UXO density, which should be a major consideration in the risk evaluation.

ORDNANCE AND EXPLOSIVES RISK IMPACT ANALYSIS

Key Features

OERIA is intended for use in assessing the relative risk reduction provided by different UXO response options. It was developed by the U.S. Army Engineering and Support Center in Huntsville. The model is intended to facilitate communications with stakeholders by eliminating, or downplaying, the use of statistical methods for UXO risk assessments. OERIA assesses explosion risks only. It does not address human health and ecological risks from munitions constituents.

The output is ordered categorical: each response action's effectiveness in reducing risk is given a letter grade, with a grade of A indicating best performance. Data requirements are minimal, and required input to the model can be obtained at very low cost. The only information required is the type of UXO present, the depth, information about types of activities that will occur after the UXO response is completed, and information about site access (whether man-made or natural barriers exist to restrict contact by the public).

Description

The model consists of a table of risk factors (see Table 3.8) that the evaluator is supposed to fill out.

First, the evaluator establishes baseline values for the factors shown in Table 3.8 prior to any response action. Then, the evaluator is supposed to establish how these values would change if different response actions were undertaken. The response action that is most effective in reducing the risk factor is given a rating for that factor of "A"; the one that is second-most effective is given a rating of "B"; and so on. Then, each response action is given an overall effectiveness rating, also on a letter scale. Table 3.9 shows an example evaluation.

Table 3.8
Input Information for OERIA

OERIA Input Factor	Possible Values for Input Factor	Explanation
Ordnance characteristics		
Type	0–3	0 = inert OE or scrap; 1 = can cause minor injury; 2 = can cause major injury; 3 = deadly
Sensitivity	0–3	0 = inert OE or scrap; 1 = unfuzed but has residual risk; 2 = less sensitive; 3 = very sensitive
Density	Use actual values	NA
Depth	Use actual values	NA
Site characteristics		
Access	Complete restriction, limited restriction, or no restriction	Complete restriction = all entry points controlled; limited restriction = man-made barriers, vegetation, water, or terrain restrict access; no restriction = no man-made or natural barriers to access
Stability	Stable, moderately stable, or unstable	Stable = OE should not be exposed by natural events; moderately stable = may be exposed; unstable = likely to be exposed
Human characteristics		
Activities	Low, moderate, or significant	Low, moderate, or significant rating is determined from matrix of activities and depth of UXO given in OERIA
Population	Estimated number of people who will use the UXO site per day	Estimate is based on demographic information

Risk Calculation Features

OERIA has a number of problems. First, it leaves too much room for value judgments that could have a major influence on the outcome. For instance, the sensitivity of the UXO is supposed to be given a value between 0 and 3, but no quantitative guidelines are provided for making the determination of whether the ordnance is "very sensitive" (category 3), "less sensitive" (category 2), "has a residual risk" (category 1), or is "inert" (category 0). In contrast, the evaluator must provide a specific measurement for UXO depth. This mix of quantitative and qualitative information is masked in the final output as a letter "grade." As a result, the model gives the appearance that it is based on quantitative inputs, while at the same time masking actual technical information. Furthermore, the results of the model are not likely to be reproducible. That is, two different evaluators could easily reach two different conclusions about which response action is likely to be most effective.

Another problem with the model is that the output it produces does not indicate the magnitude of risk reduction achieved for different response options. Thus, it is not possible, using this model, for the risk manager to determine how much extra safety can be "bought" with higher investments in response technologies. That is, one cannot determine whether one alternative, which costs more than a second alternative, actually reduces risk significantly for the additional funds invested.

A further problem is that the model makes no provisions for quantifying uncertainty. In fact, uncertainty is not even mentioned in the instructions.

Implementation Features

The instructions for using OERIA are unclear. For example, the evaluator is supposed to select a rating of 0 to 3 for type of ordnance, but no details are given about which specific types of ordnance should receive the different ratings. More important, OERIA provides no guidance to indicate how all of the factors should be combined to produce an overall risk rating (shown as the last column in Table 3.9).

Table 3.9
Example of OERIA Output

Alternatives	Ordnance				Site			Human		Overall Rank
	Type	Sensitivity	Density	Depth	Access	Stability	Activity	Population		
Baseline risk assessment	Category 1	Category 2	0.18	0–6	No restrictions to site	Site stable	Significant	~200 per day		
No action	No impact	No impact	No impact	No impact	No impact	No impact	No impact	No impact		D
Institutional controls	No impact	No impact	No impact	No impact	A	No impact	A	A		B
Surface clear, institutional controls	No impact	No impact	B	B	A	No impact	C	B		B
Clear to detectable depth, institutional controls	A	A	A	A	No impact	No impact	B	C		A

Communication Features

Although the results of the OERIA evaluation can be easily presented, the process by which the evaluation is produced cannot be easily justified to a skeptical audience. As discussed above, the model mixes qualitative information developed from personal judgments with quantitative information that can be measured directly. There is a great deal of room for adjusting the final output by changing the qualitative input factors. Thus, it is possible for different conclusions to be reached about the effectiveness of the different response actions, depending on who conducts the evaluation. The model output is likely to be disputed by concerned members of the public, and no solid technical basis is provided for defending the results.

Conclusions

OERIA resembles IR3M in many ways and has similar limitations. The logic is not always clear. Assumptions are not explained clearly. There are many opportunities to bias output based on personal preferences. There are no provisions for quantifying uncertainty. Technical instructions are vague. Finally, the process is not easy for stakeholders to follow.

NATURAL AND CULTURAL RESOURCES BANK

Key Features

Recognizing a potential conflict between UXO response activities and the legal requirement to protect natural and cultural resources, the Army Environmental Center has developed a method for assessing whether the presence of natural and cultural resources might preclude certain response alternatives (particularly those requiring extensive excavation) (Teachman and Getlein, 2002; Army Environmental Center, 2001). This method differs from others reviewed in this study in that it does not assess risks associated with UXO but instead assesses risks to natural and cultural resources from cleaning up the UXO. The output is an ordered categorical ranking of UXO sites based on information and judgments about the prevalence of natural and cultural resources at UXO sites and the quality of information available to determine that prevalence. The purpose of the

method is to identify which sites have ecosystems or cultural resources that might be irreparably harmed by UXO clearance activities.

Description

The use of military land for ordnance testing and training often results in minimal landscape modification and restricted access to these lands due to safety concerns about the presence of UXO. An ironic result of UXO contamination is that, as the NCRB method documentation states, “Army ranges contain some of the finest wildlife habitat left in the United States.” Further, reduced human intrusion into these areas may also preserve cultural sites.

Several laws protect natural and cultural resources, including the Endangered Species Act, the Clean Water Act, the Wild and Scenic Rivers Act, the National Historic Preservation Act, and the Native American Graves Protection and Repatriation Act. These requirements must be considered when selecting UXO responses. In fact, the requirement to clear UXO from a site and legal protections of cultural and natural resources may conflict at some UXO sites because of the potential need to destroy vegetation and dig up soils. Natural and cultural resources may include wetlands, threatened and endangered species (including plants and animals), critical habitat for threatened and endangered species, archeological sites, and Native American burial sites. The NCRB method is intended to complement UXO site prioritization based on risk assessments by providing a separate prioritization based on natural and cultural resource risks that might arise from clearing the UXO.

Risk Calculation Features

The NCRB assigns value scores to different categories of natural and cultural resources. The basis of the scores is a presumed association between the score and the potential of a particular resource to interfere with UXO clearance. Legal interference seems to be the greatest concern: “Rankings are weighted based on their potential for stopping or constraining training or UXO clearance.” Table 3.10 shows resources considered and the associated scores. In this case, higher scores indicate a higher predicted probability that either the resource

will be encountered or it will be of concern to regulators. Because access to UXO sites is limited due to safety concerns, some information about natural and cultural resources must be obtained by methods other than physical observation and sampling. For this reason, the NCRB method includes an additional score to assess the reliability of the information sources used. Table 3.11 shows the values that the NCRB method assigns to different data types. Lower scores reflect a presumed higher degree of confidence in the information.

The draft version of NCRB that we reviewed provides no guidance on how an assessor should choose between scores when a value range (e.g., 1–2) is indicated; presumably this decision will be based on the assessor's judgment. Other judgments are required for other scoring decisions. For example, NCRB indicates that the value score for wetlands "may vary by wetland type and region."

Table 3.10
NCRB Resource Values

Resource	Value
Natural Resources	
Threatened and endangered species critical habitat	5
EPA priority one watershed	5
Wetlands	4
Migratory birds	3
Wild and scenic rivers	3
Coastal resources	3
Highly erosive soils	1–2
Highly permeable soils	1–2
Cultural Resources	
Archeological sites	5
Sacred sites with Native American human remains and/or burial sites	5
Sacred sites without Native American human remains	4
Significant archeological sites without Native American human remains	2–3
Significant historic buildings*	2–3

*NCRB defines significant buildings as those listed or eligible for listing on the National Register of Historic Places.

Table 3.11
NCRB Information Reliability Values

Information Type	Value
Natural Resources	
Information gathered on the ground from ranges	1
Information gathered on the installation but not from ranges	2
Remotely sensed data, fine scale (i.e., aerial)	2
Remotely sensed over the range but not ground-truthed	3
Regional remotely sensed data, coarse scale (i.e., Landsat)	4
Information gathered in range ecoregion/subregion	5
Cultural Resources	
Information/artifacts collected on the ground from ranges	1
Museum collection	2
Ethnographic information	3
Paleocontext	4
Archival and published data	5

Following the value ranking of the resources and data sources, these values are summed to yield a “cumulative value” for the site. As the value relationships are juxtaposed in the resource and information rankings (significant resources score 5, “best” information sources score 1), the result is intended to provide a maximum score for a site with significant natural or cultural resources (referred to in NCRT as “resource problems”) and low confidence in the associated sources of information. Ranges of cumulative values are associated with predictions about the effect of the natural or cultural resources on UXO remediation at the site, as shown in Table 3.12.

The mathematical rationale behind the summation process is not clear. The combination of weighted scores for both observational data (the presence of natural or cultural resources) and procedural data (information sources) is problematic. The cumulative score gives no indication about the relative weights of these component values; combining them does not appear justified. The NCRB draft documentation contains an example in which the resource and information scores are presented in separate columns and summed at the end, in which case the components of the cumulative value are

Table 3.12
NCRB Cumulative Scores and Predictive Categories

Category	Cumulative Score
Response will be restricted to activities outside the natural/cultural resource area(s)	40+
Natural/cultural resources will limit response	20–39
Natural/cultural resources may limit response	12–19
Natural/cultural resources should not limit response	5–11
Natural/cultural resources will not limit response	0–4

apparent. But if NCRB is to be used as a prioritization tool, then prioritization by cumulative value depends on the subjective weighting of the components.

The NCRB documentation recognizes that it is “highly probable” that more than one natural or cultural resource may occur at a particular site, and it suggests that the summation process to determine the relative ranking of sites could be complicated by multiple occurrences of resources. The draft documentation includes a tentative method to assess multiple occurrences. It assumes that the first occurrence of a natural or cultural resource will be the most difficult to manage programmatically. Therefore, the value score for subsequent occurrences of the same resource at any one site is reduced. For threatened and endangered species, the second occurrence is assigned a value of 2 and each additional occurrence a value of 1. Multiple wetland occurrences are ranked depending on their size: if all wetlands occupy less than 50 percent of the site, then wetlands receive a total value of 4; if they occupy more than 50 percent of the site, they receive a value of 5. Highly erosive and highly permeable lands follow a similar scheme, with a value of 1 for occupying less than 50 percent of the site and a value of 2 for more than 50 percent. The second occurrence of archeological sites or sacred sites with or without Native American human remains is valued at 2; each additional occurrence is valued at 1.

The reduced ranking for multiple occurrences may be sensible from a programmatic point of view, because the intent of this tool is to identify and address situations in which the presence of natural and cultural resources may hinder UXO response. For example, if a regulatory issue arises over the presence of a single endangered species, then the presence of multiple endangered species is likely to add only marginally to the management burden. However, this method is not robust enough to portray an accurate picture of the importance and value of natural resources at the site as a whole when one considers that resources occur at different scales. For example, a 1,000-acre site containing 499 acres of a wetlands would receive a value of 4, while a 100-acre site with 51 acres of wetlands would receive a value of 5. Conversely, a 10-acre site with 1 acre of wetland would receive a value of 4—the same as the 1,000-acre site with 499 acres of wetlands. Clearly, the use of a percentage of resource relative to total site average results in a loss of information about the spatial extent of the resources.

Implementation Features

Much of the information underlying the identification of natural and cultural resources is based on already compiled sources, such as the federal listings of endangered species and wild and scenic rivers. A knowledgeable assessor must still determine the presence of these resources at a UXO site, however, and make the various subjective judgments required to evaluate information sources and assign scores.

In an attempt to prioritize sites based on the presence of natural and cultural resources, the NCRB method reduces a set of value judgments to a single value output. Even if the method were consistently applied across sites, it would still obscure underlying judgments about relative importance of sites because of its implicit use of weights and its insensitivity to the spatial extent of the resource. However, if one accepts the assumptions of resource importance in the tool, then it could be adaptable to different sites. An acceptance of these assumptions—and the underlying presumption that the greatest value of a resource lies in its propensity to hinder UXO response—is necessary to support the predictive categories that relate to the cumulative values derived. The NCRB method includes a

caveat about the utility of its output: "These cultural/natural resource rankings are very preliminary. Extensive field testing with installation resource managers is crucial to the development of a model and technique with value to the field. . . . Close work with regulators . . . is also crucial."

Communication Features

The draft NCRB documentation includes a limited discussion of a case study at Fort Wingate in which the method was tested. The authors indicate that the simplicity of the method made it very easy to communicate and that it was well received by a Native American stakeholder group. However, the reductionist nature of the method, while required for utility as a prioritization tool, may also obscure the relationships and importance of the various natural and cultural resources at a site. Point estimates such as rank scores convey little about the characteristics of a site itself without a context for comparison, and single-value outputs may convey a false sense of confidence about the characteristics of cultural and natural resources. For example, threatened and endangered species receive a value of 5, while wild and scenic rivers receive a 3; presumably this is due to a judgment that the regulatory burden of managing threatened and endangered species at a UXO site is greater than that of a scenic river, not to any opinion about the inherent and comparative values of these resources.

The NCRB documentation explicitly states that the method's intent is to catalog resources at a site to highlight those that may require significant coordination with regulators—not to assess the actual "value" or condition of the resources. This intent should be clearly communicated if the NCRB method is to be fielded widely, because different stakeholders may assign different values to resources beyond their potential to impede training or UXO response. In light of this consideration, the NCRB method is likely to be of limited utility outside of the Army risk management community unless its assumptions about the relative importance of various natural and cultural resources are made explicit and can be informed by interested stakeholders, including regulators.

Conclusions

The intent behind the NCRB method is clear and meets a need to identify UXO sites where response actions could potentially be complicated by regulatory requirements related to natural and cultural resources. This intent necessarily affects the assignment of value to different resources based on a limited set of criteria that may not be clear or acceptable to stakeholders.

The NCRB method is intended to be a technical assessment tool that produces deterministic output to aid risk managers in prioritizing sites for response. However, it also embeds value judgments, thus crossing the line into policy issues. The arithmetic manipulations in the NCRB method are not convincingly presented, but if the tool is intended only for prioritization, then this is of lesser concern if its audience understands and accepts its underlying assumptions. In this respect, the method is as much a risk management tool as an assessment methodology.

It is not clear whether a set of sites assigned NCRB scores provides a useful decision aid to program managers without further validation. A basis for comparing sites that may occur within a small range of NCRB scores is not provided (and may not be possible to provide). It is not clear that a cumulative value score of 90 is truly different from a score of, say, 50. It is likely that program managers may consider NCRB score distributions but will also require information about the components of the score, effectively decreasing the presumed benefit of the composite scoring approach. An alternative approach might be more descriptive but include a classification of the expected programmatic obstacles posed by natural and cultural resources as low, medium, or high. This would avoid a point estimate that conceals uncertainty.

SUMMARY

The reviews in this chapter make clear that none of the available risk assessment methods for UXO sites satisfies all of the criteria for a credible risk assessment that were presented in Chapter Two. Problems include lack of sufficiently sound technical models to underpin many of the calculations, lack of approaches for evaluating uncertainty, lack of sufficient opportunities for involvement of regulators

and concerned citizens in making some of the nontechnical judgments underlying the assessments, and lack of transparency. The next chapter explains how other federal agencies have successfully addressed these and related difficulties in risk assessment.

Chapter Four

OTHER FEDERAL RISK ASSESSMENT METHODS

As this report has emphasized, the task of assessing risks at UXO sites is complicated for both technical and political reasons. Technically, the difficulties stem from the presence of multiple hazards, the lack of effective technologies for locating UXO, the many sources of uncertainty associated with UXO and its environmental setting, and the uncertainties associated with future human behavior. The risk assessor must consider hazards with very different natures: explosions of ordnance items with unknown stability at discrete points in the environment, and migration of contaminants from the ordnance area to dispersed receptors. The heterogeneous distribution of UXO, with respect to both depth and area and the lack of technology capable of quantifying this distribution complicate this task. Adding further complexity are questions about which types of human behavior will occur in the future. If the site is designated as a nature preserve, for example, will people wander off the path and dig holes? Will the land use change and the site be developed in the future? Adding to these technical challenges is a lack of trust in risk assessment among many federal and state environmental regulators and concerned community groups. Many stakeholders have expressed concern that the very use of risk assessment implies that some degree of risk above zero is acceptable.

Although the technical and political challenges associated with risk assessment at UXO sites are daunting, they are not insurmountable. Many other agencies have successfully addressed similar challenges in risk assessment and prioritization. In this chapter, we present a sampling of approaches that other agencies have used to handle common problems in risk assessment. Table 4.1 summarizes the agency methods and guidelines we reviewed in developing this

Table 4.1
Federal Risk Assessment Methods Reviewed

Agency	Program	Risk Assessment Application Reviewed
DoD	Defense Environmental Restoration Program (DERP)	Relative Risk Site Evaluation (RRSE) Primer
	Department of Defense Explosives Safety Board (DDESB)	Safety Assessment for Explosive Risk (SAFER)
	Army Chemical Stockpile Disposal Program (CSDP)	Quantitative Risk Assessment (QRA)
DOE	Office of Environmental Management	Environmental Restoration Priority System (ERPS)
	Office of Civilian Radioactive Waste Management	Total System Performance Assessment (TSPA)
EPA	Office of Solid Waste and Emergency Response	Hazard Ranking System (HRS)
	Office of Solid Waste and Emergency Response	Risk Assessment Guidance for Superfund (RAGS)
	Office of Solid Waste and Emergency Response	Ecological Risk Assessment Guidance for Superfund (ERAGS)
FAA	Office of System Safety	<i>System Safety Handbook</i>
NASA	Space Program	<i>Risk Management Procedures and Guidelines and Continuous Risk Management Guidebook</i>
	Office of Nuclear Reactor Regulation	Probabilistic risk assessment
Occupational Safety and Health Administration (OSHA)	Workplace Safety	General workplace safety standards for carcinogens

chapter. We were particularly interested in determining how other agencies approach the following:

- Use risk as a consideration in setting programmatic priorities,
- Develop quantitative estimates of risk for detailed assessments at specific sites or of specific systems,
- Assess multiple risk endpoints (i.e., consider combinations of acute and chronic hazards, or combinations of human health and ecological hazards),
- Treat uncertainties in the numbers underlying the risk assessment and in the resulting output,
- Standardize the risk analysis process,
- Constructively involve stakeholders in the development and application of risk assessment and prioritization methods, and
- Engender trust in the risk assessment process and outcome.

Previous efforts to develop UXO risk assessment methods have been weak in these areas, as discussed in Chapter Three.

This chapter is not intended to provide an exhaustive review of risk assessment practices in the federal government. Rather, it illustrates how some agencies have addressed problems similar to those that the Army and the DoD have encountered in trying to implement risk assessment and risk-based prioritization for UXO sites. Many other agencies not included in Table 4.1 engage in risk assessment. In addition, the agencies listed in Table 4.1 use risk assessment in programs other than those we reviewed. We selected the methods in Table 4.1 to illustrate the benefits and pitfalls of particular risk assessment approaches. Boxes 4.1 and 4.2 briefly describe each of the methods or guidelines in Table 4.1 and why we selected them for inclusion in this chapter. These examples provide evidence that risk-based prioritization and quantitative risk assessment are supported by a long history of development and application in helping federal decisionmakers successfully manage and reduce risks.

PRIORITIZING SITES BASED ON RISK

Federal decisionmakers have long used risk assessment to help answer programmatic, scheduling, and resource distribution questions. Programmatic decisions include setting funding or regulatory priorities at a federal agency. Scheduling priorities may be used to determine the order and pace at which risk management strategies are implemented when they compete for a common set of resources. Risk assessment can also be used as one factor to guide decisions about to whom (i.e., which populations or demographic groups) or where (i.e., which sites) risk management activities are directed. Congress recently mandated that DoD develop a risk-based prioritization of UXO sites.

As described in Box 4.1, three examples of formal, risk-based prioritization systems developed by other federal agencies are the EPA Hazard Ranking System (HRS), the Defense Environmental Restoration Program (DERP) Relative Risk Site Evaluation (RRSE) Primer, and the DOE Environmental Restoration Priority System (ERPS). All three of these methods produce qualitative rankings of sites based in full or in part on risks. In all cases the rankings produced indicate relative risk values only. As is appropriate for the purpose these methods serve, there is no attempt to correlate the rankings to probabilities of adverse events occurring.

Relative Risk Site Evaluation Primer

The RRSE sorts sites in the Defense Environmental Restoration Program into three risk-related categories: “low-risk” sites, “medium-risk” sites, and “high-risk” sites. This categorization is then used as a basis for allocating program funds. As mentioned in Box 4.1, RRSE does not include provisions for evaluating risks from UXO.

The RRSE framework is based on a straightforward evaluation of contaminant sources, migration pathways, and receptors found at a particular site. It considers four environmental media and two end points:

- groundwater (human end point),
- surface water (human and ecological end points),

Box 4.1. Risk-Based Prioritization Methods

Relative Risk Site Evaluation (RRSE) Primer. The Defense Environmental Restoration Program uses the RRSE primer to sequence work and allocate the budget for environmental restoration sites across the nation (Department of Defense, 1997b). The RRSE considers only dispersed chemical contaminants in the environment, not UXO. We selected it as an example of a well-established framework for prioritizing sites for environmental restoration.

Hazard Ranking System (HRS). Under the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), Congress required that the EPA develop a “National Priorities List” (NPL) of the nation’s riskiest contaminated sites; some 1,200 sites are presently included on this list. HRS is the scoring process used to prioritize sites for CERCLA remediation (Environmental Protection Agency, 1982, 1990). We selected HRS for review because it is well established and is endorsed by the EPA for prioritizing contaminated sites based on risk.

Environmental Restoration Priority System (ERPS). ERPS was designed to help DOE decisionmakers allocate resources for environmental cleanup among 113 former nuclear weapons installations. A technical review praised the overall ERPS design, but the DOE stopped using it due to political pressures (Jenni et al., 1995; National Research Council, 1994b). We selected ERPS for review because it illustrates the potential for developing a method for prioritizing resource allocations for an extensive national environmental remediation problem while also illustrating the pitfalls to avoid in implementing such a system.

- sediments (human and ecological end points), and
- surface soils (human end point).

For each environmental medium, RRSE calculates three factors using a series of worksheets:

- the “contaminant hazard factor” (CHF), which relates concentrations of contaminants at a site to benchmark values,
- the “migration pathway factor” (MPF), which indicates the likelihood and extent of contaminant migration, and
- the “receptor factor” (RF), which estimates the likelihood of receptor exposure to contamination.

Box 4.2. Site-Specific Risk Assessment Methods and Guidelines

Safety Assessment for Explosive Risk (SAFER). The DoD Explosives Safety Board uses SAFER to evaluate the safety of proposed new sites for munitions storage (DoD Explosives Safety Board, 1999, 2000). Safe siting of new munitions storage facilities depends on establishing safe separation distances between explosive storage sites and locations where people might become injured in the event that the stored munitions explode; SAFER provides a risk-based method for making this type of evaluation. We selected SAFER as an example of a detailed, quantitative method for assessing site-specific risks associated with an acute event (in this case, explosion of a munitions storage depot).

Risk Assessment for the Tooele Chemical Agent Disposal Facility. To comply with the international Chemical Weapons Convention, as well as with U.S. Public Law 104-484, the U.S. military is in the process of destroying its stockpiles of chemical weapons. The Tooele Chemical Agent Disposal Facility (near Salt Lake City, Utah) was the first chemical weapons incinerator to become operational in the continental United States. The Army Chemical Stockpile Disposal Program developed a comprehensive assessment of the risks of the incinerator operation at Tooele as part of the process of obtaining regulatory permits to operate the facility (National Research Council, 1997b). We selected the Tooele Risk Assessment as an example of a detailed, site-specific approach for assessing risks from both acute hazards (in this case, the accidental release of a chemical weapons agent) and chronic hazards (smokestack emissions from the weapons incinerator).

Risk Assessment Guidance for Superfund (RAGS) and Ecological Risk Assessment Guidance for Superfund (ERAGS). The RAGS manuals codify the detailed, site-specific methods for assessing human health risks at contaminated sites that are being cleaned up under CERCLA (Environmental Protection Agency, 1989). The ecological counterpart to the RAGS manuals is ERAGS (Environmental Protection Agency, 1997a). We selected RAGS and ERAGS for review because they encompass widely accepted methods for quantitatively assessing risks of contaminants in the environment. RAGS, however, does not contain guidance or recommended methods for assessing explosion risks from UXO.

Probabilistic Risk Assessments for Nuclear Power Plants. Since the accident at Three Mile Island in 1979, the NRC has institutionalized the use of probabilistic risk assessment in the regulation of nuclear power plant safety (Rechard, 1999). We reviewed the probabilistic risk assessment methods of the NRC because the nuclear power industry and the NRC were pioneers in the development of such methods, which now are employed by a number of other agencies as well as private industries for assessing risks of undesired scenarios.

Box 4.2—continued

Total System Performance Assessments of Proposed Yucca Mountain Nuclear Waste Repository. The basis for the ongoing DOE assessment of risks of opening the Yucca Mountain nuclear waste repository (100 miles northwest of Las Vegas) is a method known as the “Total System Performance Assessment” (Bechtel SAIC Company, LLC, 2001). We considered the DOE’s TSPA for Yucca Mountain an example of a comprehensive, systematic analysis of risks associated with accident and contaminant release scenarios that might occur long into the future.

Federal Aviation Administration (FAA) System Safety Handbook. The FAA mandates that formal risk assessments be carried out for any decisions that “either create or could be reasonably estimated to result in a statistical increase or decrease . . . in personal injuries and/or loss of life and health, a change in property values, loss of or damage to property, costs or savings, or other economic impacts valued at \$100,000,000 or more per annum” (FAA, 1998). Different programs within FAA use different risk assessment methods, but they commonly employ probabilistic risk assessment techniques such as those used by the NRC (FAA, 2000). We included a review of FAA’s *System Safety Handbook*, which spells out FAA doctrine on risk assessment, as an example of the established use of formal risk assessment to manage the potential for low-probability but high-consequence events (such as an airline crash) (FAA, 2000).

National Aeronautics and Space Administration (NASA) Risk Management Procedures and Guidelines. Because NASA engages in high-risk, time-sensitive operations for which often only one trial is possible, quantitative risk analyses are central to its programs. NASA’s risk assessment protocols are codified in its *Risk Management Procedures and Guidelines* (NASA, 2002c). The agency’s risk assessment protocols are based on probabilistic risk assessment tools such as those used by the NRC (see NASA, 2002a,b). We included a review of NASA risk assessment guidelines as another example of the institutionalized use of formal risk assessment in making tradeoffs that might affect the likelihood of high-consequence, risky events.

Occupational Safety and Health Administration (OSHA) Standards for Workplace Safety. OSHA has well-structured processes for evaluating risks of occupational exposure to contaminants in the workplace. We reviewed general OSHA occupational exposure risk assessment methods as a potentially relevant model for quantitative, site-specific risk assessment.

According to the RRSE documentation, “For each medium, factor ratings are combined to determine a medium-specific rating of high, medium, or low. A site is then placed in an overall category of high, medium, or low based on the highest medium-specific rating.” For example, if a site’s overall groundwater rating is high but ratings for surface water, sediments, and surface soils are all low, the site would receive a “high” classification. Figure 4.1 shows the relationships among the media, evaluation factors, media-specific relative risk rating, and the overall site category.

For a single contaminant, the CHF is computed from the ratio of the maximum concentration of that contaminant to a reference value. For carcinogens, the reference value is the concentration that represents a 10^{-4} risk of increased cancer incidence. For noncarcinogens,

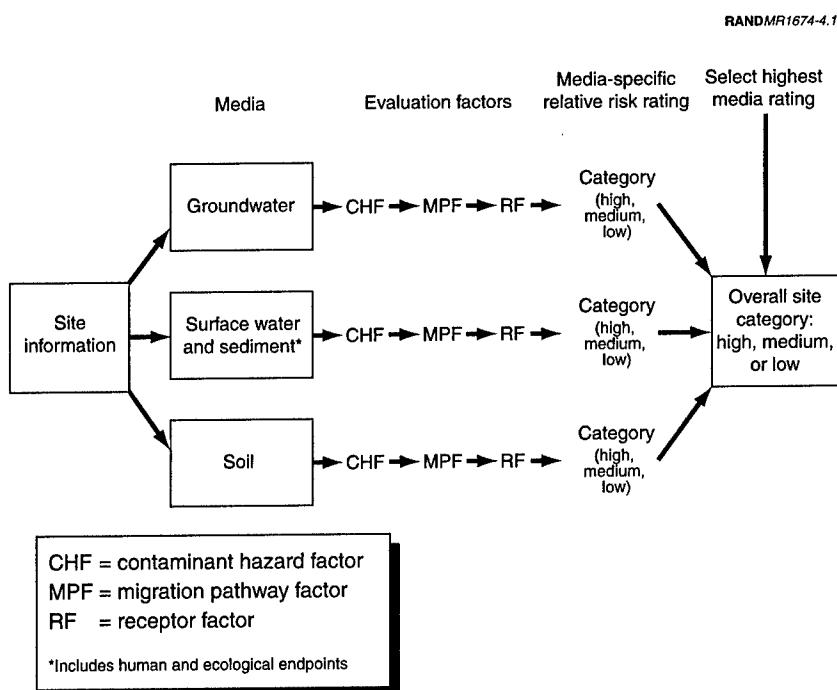


Figure 4.1—Flow Diagram Showing Inputs to and Output from the Relative Risk Site Evaluation

the reference value is the “daily reference dose,” below which adverse noncancer health effects are unlikely to occur. When more than one contaminant is present (at any concentration above analytical detection limits), then the CHF is determined from the sum of the ratios described above. Table 4.2 shows how the CHF is classified.

The technical basis for the breakpoint values shown in Table 4.2 is not related to risk in any absolute sense. Rather, a DoD working group looked at data from “thousands” of DoD sites, calculated ratios, and then sought breakpoint values that would evenly distribute the ratings of significant, moderate, and minimal among sites. Overall, the classification of the contaminant hazard factor is a “worst-case” characterization of contaminant hazard as opposed to a realistic assessment of actual hazard at a given site.

The MPF is classified as “evident,” “potential,” or “confined.” Criteria for classification are specific to each environmental medium. As an example, Table 4.3 shows the RRSE guidance for determining the MPF for groundwater.

The technical basis for this classification rests on the quality of existing data and professional judgment used to infer migration. Distance from the source, specifically at a point of exposure, is used as the evaluation criterion. This measure does not require judgments to be made about groundwater velocities or contaminant travel times in groundwater. No distinction is apparently made about which contaminant might have migrated the farthest. Thus, a relatively low-risk but mobile contaminant could cause the MPF to be “evident,” when in fact a high-risk contaminant could be less mobile.

Table 4.2
Calculation of CHF

Sum of Ratios of Maximum Contaminant Concentrations to Reference Values	CHF “Value”
> 100	Significant
> 2 and \leq 100	Moderate
\leq 2	Minimal

Table 4.3
Calculation of MPF for Groundwater

MPF	Criteria
Evident	Analytical data or observable evidence indicates that contamination in the groundwater is moving or has moved away from the source area
Potential	Contamination in the groundwater has moved only slightly beyond the source (i.e., tens of feet), could move but is not moving appreciably, or information is not sufficient to make a determination of "evident" or "confined"
Confined	Information indicates that the potential for contaminant migration from the source via the groundwater is limited (due to geological structures or physical controls)

The default assumption for the MPF in the absence of sufficient data is "potential," and thus the RRSE is conservative in this feature. For groundwater, the application of the "confined" classification is based on the existence of physical barriers to migration, low precipitation, or long travel times from the source of contamination to receptors. Because the RRSE is only to be used at sites where some level of site characterization has already been completed, the implicit assumption is that sufficient information is available to make this more technically demanding determination of contaminant travel times. The framework requires the user to apply professional judgment in determining whether sufficient data exist.

The RF is classified as "identified," "potential," or "limited" based on available information, detailed instructions provided in the RRSE Primer, and professional judgment.

As shown in Figure 4.2, the RRSE combines the CHF, MPF, and RF for each medium. Then, as noted previously and indicated in Figure 4.1, the overall classification for a site is the highest classification among all the relevant media at the site.

The RRSE Primer makes clear that the "grouping of sites into high, medium, or low relative risk categories is not a substitute for either a baseline risk assessment or health assessment; it is not a means of placing sites into a Response Complete/No Further Action category; and it is not a tool for justifying a particular type of action (e.g., the

RANDMR1674-4.2

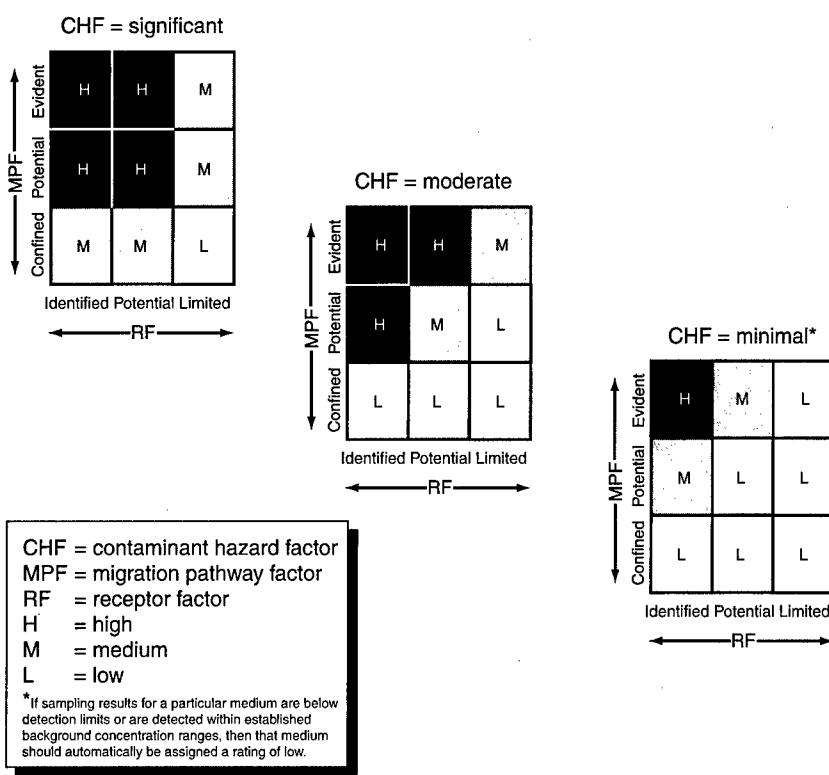


Figure 4.2—RRSE Matrix for Assigning Site Scores

selection of a remedy” (emphasis in original). The primer also states, “Like any risk evaluation tool and perhaps more so than a comprehensive risk assessment, the relative risk evaluation framework makes use of assumptions and approximations.”

While RRSE does not address explosion risks, it is a good example of how risks can be considered (in this case, qualitatively) in setting priorities among sites. It does not purport to measure risk in any absolute sense. Indeed, because of the inherent conservatism of its hazard, migration, and exposure factors, RRSE can at best be viewed as a

very coarse tool for programmatic planning, budgeting, and priority setting for implementation.

The RRSE highlights two limitations of qualitative methods and how they are commonly addressed. First, qualitative methods rely on ordinal data, as opposed to interval or ratio data. For example, the RRSE method sorts hazards into several bins and combines them to form an aggregate score. Some argue that it is inappropriate to use mathematical operations (e.g., addition and multiplication) on such ordinal data. However, others suggest that a metric should be judged more by its usefulness than by its adherence to a set of scale-based rules (Dawes, 1994; Velleman and Wilkinson, 1993). A high school student's grade-point average is a common example of a useful metric that violates scale-based rules. The successful application of the RRSE method lends some level of support to the qualitative metrics produced.

Second, the categorization or grouping required by the RRSE method introduces a great deal of subjectivity into the assessment process. Subjectivity can limit the ability of a risk manager to make comparisons among assessments conducted at different sites or by different analysts. The RRSE method attempts to address this limitation by providing specific definitions for each category. If category definitions are able to eliminate individual judgments from the assessment process, risk managers can be more confident of making comparisons across assessments.

Hazard Ranking System

Like the RRSE framework, the HRS illustrates how risk assessment is currently used in practice to prioritize contaminated sites for remediation (Environmental Protection Agency, 1982; Environmental Protection Agency, 1990). The HRS is a structured value analysis resulting in a single numerical score between 0 and 100. Higher values are intended to indicate a greater perceived degree of hazard. A site must have a score of at least 28.5 to be eligible for inclusion on the list of sites that are eligible for cleanup under CERCLA.

Similar to the RRSE, the HRS score is derived from pathway-specific scores for the following four pathways:

- groundwater migration,
- surface water migration,
- soil exposure, and
- air migration.

If there are multiple affected groundwater systems or watersheds at a site, pathway scores are calculated separately for each entity, and then the maximum is chosen for use in calculating the overall HRS score.

The score for each pathway considers human health risks and ecological risks. As a general matter, these two broad risk classes (human health and ecological) are combined into a single metric in the calculation of the four pathway-specific scores and in the overall HRS score. HRS prescribes in detail how to perform this combination. In all cases, the total contribution to the pathway score due to ecological risks is limited to 60 of the maximum total of 100 points.

The composite HRS score is derived from the four pathway scores (each ranging from 0 to 100) as the square root of the mean of the squared pathway scores. Each of these four pathway scores is in turn based on consideration of one or more “threats.” The groundwater and air migration pathways each consist of a single threat. The surface water migration pathway considers three different threats (drinking water, human food chain, and environmental), while the soil exposure pathway considers two threats (resident population and nearby population). In the pathways that consider multiple threats, the individual threat scores are scaled and summed to ensure that the final pathway score is between 0 and 100.

In addition to the general combination of human health and ecological risks, the HRS relies on several other notions of combining risks. These include, among others, the following:

- combining risks from different sources within a site,
- combining risks from different hazardous substances in a given pathway and threat, and
- combining risks to different entities within a pathway (e.g., watersheds and aquifers).

HRS prescribes how to perform these sorts of combinations under different circumstances.

It is important to emphasize that like the RRSE output, the HRS score does not measure risk but is intended to provide a relative ranking of hazardous potential for different sites. For example, according to EPA, “Because the HRS is intended to be a screening system, the Agency has never attached significance to the cutoff score [of 28.5] as an indicator of a specific level of risk from a site, nor has the Agency intended the cutoff to reflect a point below which no risk was present” (Environmental Protection Agency, 1992). Thus, although higher HRS scores are intended to reflect higher levels of risk, knowing only the HRS score without further information provides only a coarse indication of the severity of the risks and no indication of the nature of the risks.

The HRS data requirements are far more extensive than those for the RRSE process. Nonetheless, because HRS is a screening tool, it is governed by a general philosophy that suggests using easily available sources of information first and performing more invasive data collection only if it is likely to affect the HRS score.

The HRS shows how qualitative scoring definitions that are more detailed than those of the RRSE can be used successfully. The instructions about how to calculate the HRS scores are explicitly defined, and the documentation does a reasonably good job of anticipating the myriad of site conditions that scorers might face. This helps to eliminate subjectivity on the part of the implementer. Also, the HRS stresses the importance of clearly documenting how scores were determined (e.g., what data were used). Moreover, the system is designed to be robust against some of the more uncertain inputs into the score, such as the number of affected individuals or the amounts of hazardous waste quantities. Thus, in most cases, if two people can agree about the order of magnitude of these quantities, then they will agree on that component of the score.

A significant limitation of the HRS is its integration of human health and ecological risks into a single score. While scientifically sound technical models underlie the HRS score (such as those for estimating migration, bioaccumulation, and toxicity), the integrity of their outputs is effectively compromised in the composite HRS score. This

is because the HRS uses a complicated sequence of operations on key scientific inputs to produce the final score, with little or no justification for the manner in which these elements are combined. The scores that certain elements are assigned as well as how these scores are aggregated (e.g., add, multiply, choose the maximum or minimum, take the mean, take the geometric mean, etc.) are largely subjective. The weightings of elements and their contributions to the score are endogenous to the scoring process and reflect value judgments that may or may not be commensurate with those of stakeholders.

Environmental Restoration Priority System

One way to avoid the problems identified in the HRS is to be sure that the values assigned to different risk types or risk-reduction objectives are explicit in the scoring process. The DOE's ERPS system provided for such an explicit statement of relative values. DOE developed ERPS in the early 1990s "to allocate [appropriated funds] among field offices, programs, and installations" (DOE, 1991). DOE cancelled the use of ERPS after the initial implementations because of conflicts with stakeholders (described later in the chapter). Nonetheless, ERPS received relatively favorable technical reviews, and it illustrates a possible approach for integrating multiple objectives in a prioritization process (Jenni et al., 1995; National Research Council, 1994b).

The ERPS design was based on a formalized, mathematical concept known as "multiattribute utility" theory. Multiattribute utility theory was developed to evaluate how well different alternative actions satisfy multiple, predetermined objectives. The approach involves developing "subscores" for each objective and then using systematic techniques to aggregate these subscores. A National Research Council evaluation of methods for ranking hazardous waste sites for remedial action cites two texts that describe the theoretical basis for this approach: Keeney and Raiffa (1976) and Edwards and Newman (1982) (National Research Council, 1994b).

ERPS evaluated the ability of individual packages of potential cleanup actions known as "budget cases" to achieve six program objectives. The program objectives, drawn from the overall goals of

DOE's Environmental Management Program, were to accomplish the following:

1. reduce health risks to workers and the public,
2. decrease adverse environmental impacts,
3. lessen adverse socioeconomic impacts,
4. comply with regulatory requirements,
5. decrease long-term cleanup costs (e.g., by cleaning up a spreading problem quickly), and
6. decrease uncertainties related to risks and costs (Jenni et al., 1995; National Research Council, 1994b).

Each DOE installation developed three to ten budget cases to feed into ERPS. Each budget case described the total set of remediation activities that the installation could carry out if it were to receive funding sufficient to cover the costs of all the activities included in the case. Each installation developed a "maximum case" consisting of all the activities the facility could manage with unlimited funding, a "minimum case" consisting of activities possible with a prescribed minimum level of funding (usually 70 percent of the prior year's budget), and one or more intermediate cases with activities that could be carried out at a total cost in between the maximum and minimum case costs (Jenni et al., 1995). DOE installation managers were trained in the use of ERPS and were responsible for developing an ERPS score for each of their proposed budget cases.

As noted above, the total ERPS score was based on a summation of "subscores" indicating the potential for each budget case to achieve each of the six Environmental Management Program objectives. The ERPS equation for combining the subscores was as follows (National Research Council, 1994b):

Utility of budget case is

$$W_{pr}U_{pr}(S_{pr}, S_u) + W_{ir}U_{ir}(S_{ir}, S_u) + W_{env}U_{env}(S_{env}) + W_{soc}U_{soc}(S_{soc}) \\ + W_{ur}U_{ur}(S_{ur}) + W_{rr}U_{rr}(S_{rr}),$$

where

W = relative "weight" assigned to the objective,
 U = utility function that transforms "scores" for each objective
 into a measure of value,
 S = score for the particular objective,
 pr = health risk to the surrounding population,
 ir = health risk to an individual,
 u = risk urgency,
 env = environmental risks,
 soc = socioeconomic risks,
 ur = uncertainty reduction,
 rr = regulatory responsiveness.

ERPS developers established the weights for each objective (the "W" values in the above equation) through a formal elicitation. The elicitation process involved surveying Environmental Management Program officials about how much they would be willing to pay for progress toward one objective at the expense of progress toward another objective. Table 4.4 shows the weights assigned to each objec-

Table 4.4
Weights Assigned to the Objectives of the DOE Environmental Management Program

Objective	Weight
Health risk reduction	36
Environmental risk reduction	13
Socioeconomic impact reduction	9.5
Regulatory responsiveness	9.5
Uncertainty reduction	32
TOTAL	100

SOURCE: Jenni et al. (1995).

tive. Intuitively, these weights represent the importance, from the DOE managers' perspective, of each objective relative to the other objectives. Thus, health risk reduction, with the highest weight, was most important, followed by uncertainty reduction.

ERPS included detailed protocols for calculating each subscore (the "S" values in the equation) and for each utility function. These methods were different for every objective. In general, the scores were on a 1–7 scale, with 1 indicating favorable conditions and 7 indicating very unfavorable conditions (National Research Council, 1994b). The utility functions (represented by the "U's" in the equation) were generally exponential (National Research Council, 1994b).

Both the National Research Council and a specially appointed commission known as the "Technical Review Group" evaluated the technical merits of ERPS. The Technical Review Group concluded that

This methodology represents the state of the art. However, it has major limitations in what it can accomplish even with perfect input. Even with its current limitations, the system can play an important role in ordering priorities (Burke et al., 1991, as quoted in National Research Council, 1994b).

The National Research Council said that it could not reach a conclusion about the ultimate usefulness of ERPS, based on the information it had received, in part because ERPS had not been fully implemented at the time of its review. However, the Research Council pointed out several limitations. One of the most important was the frequent use of subjective judgment as input when data were lacking:

ERPS allows input of subjective judgment in the scoring if the scorer does not have adequate scientific data available. This does lead to user-friendliness, but also could be subject to abuse and poor decisionmaking. The tradeoff here between user-friendliness and sound ranking of sites needs consideration (National Research Council, 1994b).

Thus, although the weights assigned to the different objectives (as shown in Table 4.4) were clearly defined, many of the inputs used to determine the subscores were not so well defined.

The multiattribute utility approach to combining subscores for individual objectives might be useful as a mechanism for prioritizing sites based on explosion risks and munitions constituents risks. However, care would need to be taken to ensure that the prioritization system was not subject to abuse based on manipulation of subjective inputs. Further, because explosion risks seem the most time critical at UXO sites in that they pose acute risks and prevent further cleanup or redevelopment activity, it would seem wise to adopt a prioritization system that considers explosion risks first, rather than trying to assign relative weights to explosion and constituents risks.

DEVELOPING QUANTITATIVE RISK ESTIMATES FOR SPECIFIC SITES

Risk-based prioritization methods can approximate the relative risk of one site versus another. However, their output is not sufficiently detailed to support decisions at specific sites because the correlation between the outputs and the probability of adverse events actually occurring is very weak at best. Detailed, quantitative methods for assessing the probability of adverse events have evolved in many federal agencies over the past several decades.

The process of performing these assessments generally follows the principles outlined by the National Research Council and described in Chapter Two. However, differences in the nature of the involved risks as well as legislative and societal events have produced a variety of methods and tools for carrying out site-specific, quantitative risk assessments (National Research Council, 1996). For example, health risk assessments conducted under EPA waste-site programs focus on analyzing scenarios in which contaminants migrate through the environment to humans. The quantitative techniques developed for this purpose include environmental models for predicting the concentrations of contaminants that might be transported through water, air, or soil as well as dose-response models for estimating probabilities of health effects based on the predicted contaminant concentrations in the environment. In contrast, while risk assessments to comply with NRC requirements also must consider the potential for environmental transport of and human exposure to contaminants, much of the basis for risk assessments in the nuclear power industry has evolved from systems engineering (Haimes,

1998). A nuclear power plant risk assessment focuses on identifying potential power plant failure scenarios and determining the probability of those scenarios occurring. Rechard (1999) notes that there is “a subtle difference between risk assessments for hazardous chemicals and those of nuclear facilities in that assessments for hazardous chemicals have a less intimate connection to systems (engineering) analysis.”

For the UXO problem, both the environmental exposure and systems approaches to risk assessment may be useful. Examples of the former are the EPA processes for assessing human health and ecosystem risks at CERCLA sites as described in a series of publications known as the *Risk Assessment Guidance for Superfund* (RAGS) manuals and in the *Ecological Risk Assessment Guidance for Superfund* (ERAGS) (EPA, 1989, 1997a). An example of the latter is the NRC’s use of probabilistic risk assessment for evaluating the safety of nuclear power plants. Another method is the DoD Explosives Safety Board’s *Safety Assessment for Explosion Risks* (SAFER), which represents an approach to system failure assessment that is less formalized than the NRC process.

Risk Assessment Guidance for Superfund

The RAGS manuals (there are four) explain how to conduct risk assessments at CERCLA sites in accordance with EPA requirements (EPA, 1989). RAGS details how to (1) establish baseline (i.e., existing) risks, (2) estimate levels of chemicals that can remain while adequately protecting human health, and (3) compare risk reductions for alternative remedial actions. RAGS Part A describes how to conduct a baseline risk assessment. Part B focuses on the establishment and refinement of “preliminary remediation goals,” which in effect define what the desired outcomes (in terms of achievement of certain risk levels) ought to be for the site remediation process. Part C addresses the formal comparison of remedial response alternatives with regard to both chronic and acute human health risks. RAGS Part D differs somewhat from the other parts in that it revisits and refines various issues addressed in Parts A through C, suggesting improved methods and guidance.

The RAGS process is similar in spirit to that described by the National Research Council (National Research Council, 1983). Similar to the

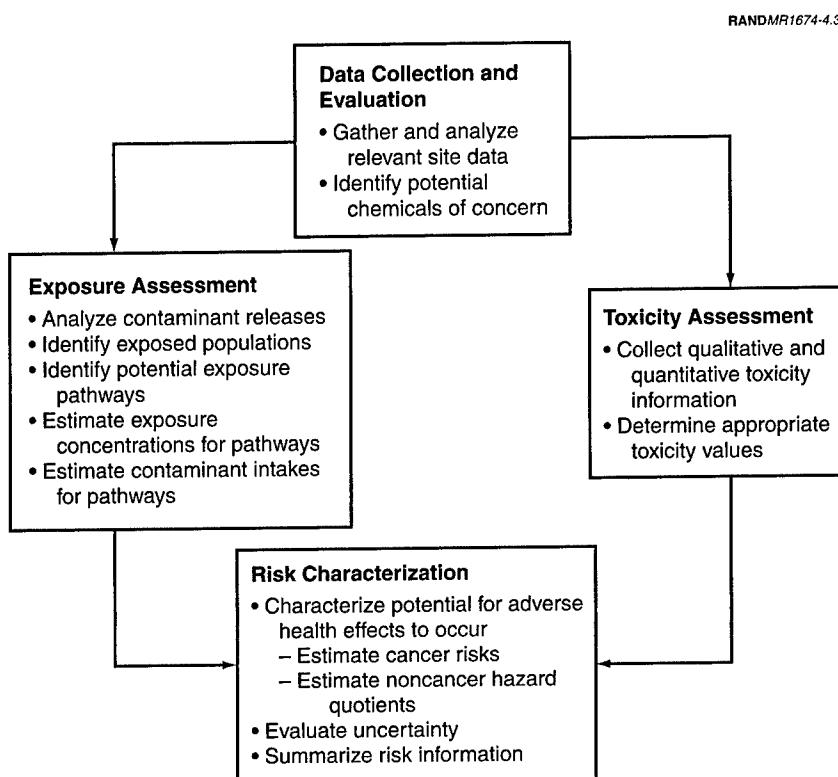
National Research Council process, a RAGS evaluation has the following four steps (see Figure 4.3):

1. **Data collection and evaluation.** Identify substances present, and gather data necessary to evaluate risks.
2. **Exposure assessment.** Estimate the magnitude of actual and/or potential human exposures.
3. **Toxicity assessment.** Identify whether exposure to the contaminant of concern can increase the incidence of an adverse health effect (such as cancer or birth defects), and quantitatively evaluate the relationship between the amount of contaminant received and the incidence of adverse health effects in the exposed population.
4. **Risk characterization.** Summarize and combine the exposure and toxicity assessments to characterize baseline risk, “both in quantitative expressions and qualitative statements.”

The exposure-quantification step begins by assessing all possible routes by which humans might be exposed to contamination (see Figure 4.4). Exposure concentrations are based either on environmental samples gathered from potential points where humans might be exposed or on complex mathematical models that predict those concentrations based on information about contaminant source areas and the environmental medium of concern (air, water, or soil). Estimates of exposure are developed for both current and future land-use assumptions.

For the toxicity assessment, the RAGS manuals indicate that the risk assessor should rely on previously compiled information about contaminants. For example, the Agency for Toxic Substances and Disease Registry maintains a database of contaminants found at CERCLA sites; the database contains comprehensive surveys of existing health studies on the effects of those contaminants. Thus this step does not involve gathering new site-specific information, but compiling existing data on the health effects of the contaminants of concern.

The risk-characterization step is based on the “reasonable maximum exposure,” rather than on a range of possible exposures. RAGS defines reasonable maximum exposure as “the highest exposure that is

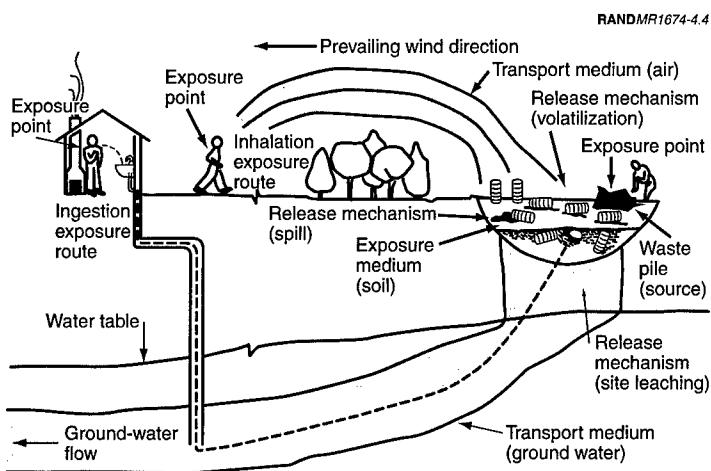


SOURCE: Adapted from EPA (1989).

Figure 4.3—RAGS Risk Assessment Steps

reasonably expected to occur at a site for each pathway and for the sum of all pathways.” RAGS recommends using as the reasonable maximum exposure the 95 percent upper confidence limit on the arithmetic average of the exposure concentrations experienced over the exposure period.

Although it is impossible to anticipate all possible site conditions, the RAGS processes are designed to be general enough to apply at virtually all CERCLA sites. In addition, RAGS Part D states that the quantitative risk assessment methods developed therein would be applicable to “non-NPL, BRAC and Brownfield” sites, as well as to



SOURCE: Adapted from EPA (1989).

NOTE: Surface water and food-borne exposure pathways, though addressed in RAGS, are not illustrated here.

Figure 4.4—Example Pathways for Exposure to Contaminants from a Waste Site, as Illustrated in RAGS

CERCLA sites. RAGS also provides specific recommendations for selected nonstandard conditions (e.g., radioactive wastes), further specifying that when other nonstandard situations are encountered, expert judgment should be used to adapt the standard guidelines.

The contamination scenarios addressed by RAGS are in principle similar to those resulting from munitions constituents that might disperse in water or soil. In our view, RAGS appears to be a useful approach for performing a technical evaluation of risks to human health from munitions constituents dispersed in soil and water. However, RAGS does not explicitly address explosion risks from UXO.¹ RAGS acknowledges that there may be “hot spots” of contamination with potential for acute exposures but does not provide guid-

¹If CERCLA hazardous wastes present an explosion risk, such as from gaseous emissions, this would be addressed, although probably as a CERCLA emergency response.

ance on how to evaluate the likelihood of exposure at a hot spot, especially a hot spot such as a buried UXO item that cannot be located with certainty. Thus the RAGS manuals, on their own, are not sufficient for assessing risks present at UXO sites; supplementary methods are needed to assess explosion risks. Having a credible quantitative explosion risk assessment method is especially important once the UXO clearance process is complete because, as Chapter One explains, no detection technology can guarantee that every UXO item has been located and removed.

Ecological Risk Assessment Guidance for Superfund

ERAGS is the ecosystem risk assessment counterpart to RAGS. The nature of the ERAGS technical guidance and evaluation process is very different from that specified in RAGS. For example, the numerous technical details necessary for performing the risk calculations are for the most part not addressed by ERAGS. The EPA provides guidance on many of these technical details in other agencywide documents on ecological risk assessment (EPA, 1997a).

Eight steps form the basis for the formal ERAGS process:

1. screening-level problem formulation and ecological effects evaluation,
2. screening-level preliminary exposure estimation and risk calculation,
3. baseline risk assessment problem formulation,
4. data quality objectives and study design,
5. field verification of sampling design,
6. site investigation and analysis of exposure and effects,
7. risk characterization, and
8. risk management.

The first two screening-level steps establish basic site characteristics, such as locations of sensitive habitats, types of wildlife present, types of chemicals present and upper bounds on their concentrations, and probable key pathways to the affected end points. The key end

points of the screening-level steps are estimated hazard quotients for the various chemical-pathway/end-point combinations perceived to be of most concern. Hazard quotients represent the ratio of the estimated dosage to the dosage believed to result in no adverse effects. However, because these quotients are based on screening-level data only, they are designed to be conservative, effectively delimiting worst-case ecological risk scenarios. This is in accordance with the primary purpose of these stages, which is to decide whether ecological risks are potentially high enough to warrant the detailed ecological risk assessment.

At the end of the screening stages, the risk assessor determines whether a full ecological risk assessment (stages 3–7) is necessary. Stage 3 is a far more detailed version of stage 1 and concludes with a decision point at which all parties agree on the assessment end points, the exposure pathways, the risk questions, and the site conceptual model. These form the basis for the detailed design of the study in stage 4, which includes what and how much data will be collected and how it will be analyzed. ERAGS specifies that these plans should be coordinated with the similar efforts necessary for the human health risk assessment specified in RAGS. Stage 5 examines the feasibility of the sampling plan, ascertaining whether it is actually possible to collect data of the indicated nature and specificity. In this way, any practical shortcomings can be resolved before committing to the full investment of resources, and it is not until after this stage that the final work and data-gathering plans are issued and formalized. In stage 6, all data collection and analysis are performed. Stage 7 presents the risk characterization results, including a detailed analysis of uncertainty. In the final stage, remedial decisions are made and formalized. Results from the ecological risk assessment are only part of this decision.

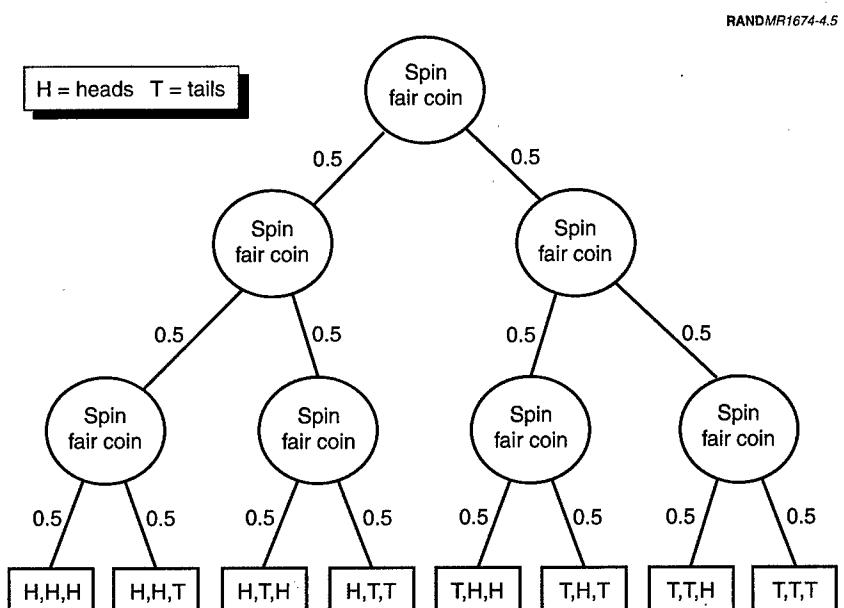
Although ERAGS might be useful for addressing ecological risks from munitions constituents, the UXO problem might warrant specialized considerations of ecological risks incurred from response actions. Such actions (e.g., vegetation clearance) have the potential to be invasive and ecologically disruptive. While it is true that some CERCLA response actions can be similarly invasive (e.g., sometimes massive amounts of soil are excavated and incinerated at CERCLA sites, with total clearance of vegetation and replacement with an impermeable cap), EPA guidance on formal assessment of ecological risks derived

from physical alterations to the environment appear to be less rigorously developed than the more traditional analysis of chemical risks. Because of the inherently invasive nature of UXO response, ecological risk assessments of UXO sites should evaluate environmental risks arising from soil excavation and other physical environmental changes that occur during UXO removal.

Probabilistic Risk Assessment (NRC, NASA, and FAA)

The NRC, NASA, and the FAA all are involved in regulating and overseeing the safety of systems that are subject to acute failures (such as nuclear power plant meltdowns, shuttle explosions, and airline crashes). All three of these agencies use a category of techniques known as “probabilistic risk assessment” (PRA) to help evaluate the risks of such failures occurring (FAA, 2000; NASA, 2002a,b,c; NRC, 1983, 1995, 1998a,b). As mentioned in Chapter Two, PRA techniques originated at Bell Labs in the early 1960s; they were used initially to help identify failures in and improve the launch success rate of Minuteman missiles. The nuclear power industry and the NRC led the development of significant advances in PRA theory and techniques beginning in the early 1970s. Since the binary nature of technological failure events (i.e., occur or not occur) is conceptually similar to UXO explosion hazards, the PRA tools used in NRC, NASA, and FAA risk assessments could provide a useful structure for UXO explosion risk assessment.

A traditional PRA begins by asking the questions “What can happen?” and “How can it happen?” Analysts then use “tree” diagrams to represent the events that could unfold after some undesired initiating event occurs. Such tree diagrams are known as “event trees” (NASA, 2002b; Shafer, 1996; Rasmussen, 1981). If the probabilities of each event in the sequence are known, then the probability of each scenario resulting from the initiating event can be determined. Figures 4.5 and 4.6 show simplistic examples of event trees. Figure 4.5 shows the possible results after three spins of a fair coin. Because each event in the sequence is independent, probabilities of each event can be multiplied to determine the chance of each final outcome. For example, the probability of spinning three heads is $0.5 \times 0.5 \times 0.5$. Figure 4.6 shows an example event tree for analyzing the risks of an uncontrolled fire in an industrial facility containing



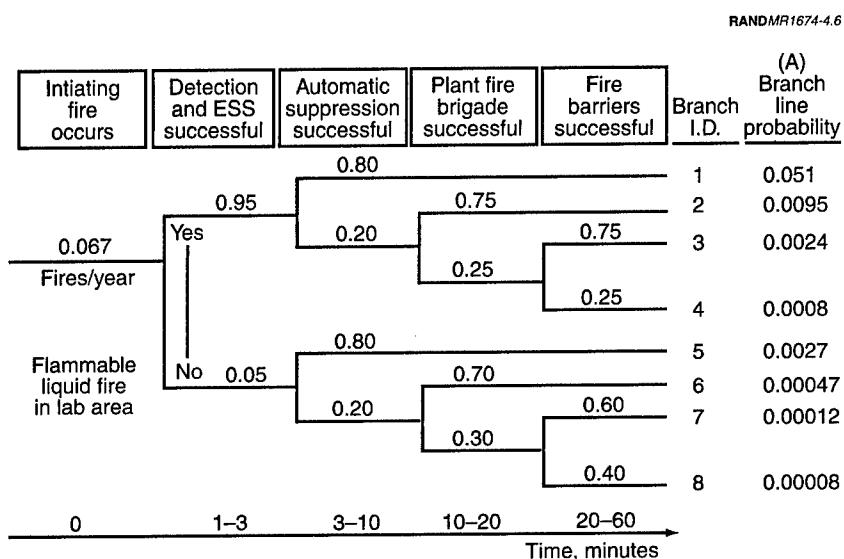
SOURCE: Adapted from Shafer (1996).

Figure 4.5—Event Tree Showing Possible Scenarios After Three Spins of a Fair Coin

flammable liquids. The event tree shows all possible scenarios that could unfold after a vessel of flammable materials catches on fire and the chance of each scenario occurring.

“Fault trees” also are used for PRA. Creation of a fault-tree diagram begins by defining a top-level, undesirable event (e.g., a specific form of system failure, for example a containment system failure in a nuclear power plant). Then, the risk assessor develops branches of the tree to characterize all sequences of events that could lead to the top-level event.

We can imagine a simple fault tree for a hypothetical situation involving loss of water pressure in a tall building. Water pressure loss occurs either if a water shortage occurs or if building pumps fail. Water shortage will occur if there is a loss of public water and the water level in reserve tanks is low. If the probabilities of all events



SOURCE: Adapted from Barry (2002).

Figure 4.6—Event Tree Showing Possible Scenarios After Flammable Liquid in Industrial Facility Catches Fire

leading to a failure event are known, the risk of the event can be calculated. For example, if all events in our hypothetical situation are independent, the risk of losing water pressure is $P_1 + P_2$. Similarly, the risk of a water shortage (i.e., P_1) would equal $P_3 \cdot P_4$. Such events are actually often dependent. Thus, these example calculations would represent an upper bound for cases where only one of several events must occur (i.e., A, B, or C must occur) and a lower bound if all events must occur (i.e., A, B, and C must occur) (Melchers, 1992).

In principle, these techniques characterize all possible system outcomes, which as a byproduct delineates all possible pathways to system failures. Together, these types of PRA tools support risk management decisions in FAA, NASA, and the NRC by enumerating possible system failures and elucidating possible risk-mitigation strategies by indicating failure pathways. However, early applications of PRA engendered criticism. For example, critics of the *Reactor Safety*

Study (NRC, 1975)—a landmark report that used fault and event trees to calculate the human health risks from operating two nuclear reactors—highlighted two of its limitations (Lewis et al., 1978; Kamins, 1975). First, they suggested, fault tree analysis and event tree analysis require quantification of the probability of each specified event. Since extensive data on human errors or equipment failure may not be available, these probability estimates may need to reflect subjective judgments. Second, the resulting risk estimates may be inaccurate if the analysis does not comprehensively account for all possible failure events and sequences, as well as their correlations and interactions. Overcoming these limitations requires clear identification of subjective judgments, so that these can be scrutinized for their legitimacy, and careful review of whether the analysis comprehensively describes the potential risk pathways.

Despite these limitations, PRA has matured, and its use has increased substantially. Following the accident at the Three Mile Island nuclear power plant in March 1979, two reviews recommended greater use of PRA and further development of tree-based tools (President's Commission on the Accident at Three Mile Island, 1979; Rogovin and Frampton, 1980). The publication of the *PRA Procedures Guide* (NRC, 1983) represented an important step in the maturation of PRA methods. Similarly, the use of PRA in NASA has grown, despite early setbacks. The use of PRA in NASA dates back to the Apollo mission in the early 1960s. History proved the empirical success rate of the Apollo (6 out of 7, or 86 percent) to be considerably higher than the 20 percent chance of success predicted by PRA, causing PRA to fall into disuse at NASA. Interest was rekindled in the aftermath of the Challenger explosion, however, with an official endorsement of the technique by the Slay Committee in 1988. NASA has since begun building a core competency in PRA theory and practice and is aggressively pursuing agencywide acceptance and use of the technique. The agency has conducted awareness and methodological training and recently released official PRA procedures guidelines (NASA, 2002a,b). The existing standard software for PRA is a program called “Systems Analysis Programs for Hands-on Integrated Reliability Evaluation” (SAPHIRE), originally developed by the NRC in the late 1980s. NASA is developing its own tool for PRA, known as the “Quantitative Risk Assessment System” (QRAS), which requires further beta testing and training before full deployment. Both SAPHIRE

and QRAS gave comparable results in a recent trial of PRA for a space station subsystem.

Safety Assessment for Explosion Risk

The Department of Defense Explosives Safety Board (DDESB) developed the SAFER tool (DDESB, 2000) to help manage risks involving explosives and munitions at existing DoD facilities, those undergoing realignment or closure, FUDS, and munitions-related construction. The primary use of SAFER is to supplement quantity-distance criteria during the siting of new explosive-storage facilities, most notably when the quantity-distance criteria cannot be met. In these cases, a “risk-based siting” may be pursued if it can be demonstrated that the explosive risks posed by the site would be acceptably low. SAFER provides model-based estimates of these risks based on site-specific information and can also be used to perform risk-based comparisons of different siting options.

SAFER estimates annual expected fatalities and individual probability of fatalities from explosion hazards posed by a “potential explosive site” (PES) (e.g., a new munitions stockpile) to surrounding “exposed sites” (ESs). The fundamental measure of risk assigned to a particular PES is the “annual expected fatalities,” defined as the average number of fatalities expected per year due to an explosion at that site. In addition to expected fatalities, SAFER considers the “individual probability of fatality.” This is the marginal (i.e., unconditional) probability of an individual dying from an explosion in a given year, accounting for both the probability of explosion and the conditional probability of fatality given an explosion.

According to the DDESB, the probability of an explosion at the PES in a single year is the most uncertain input to the model because it depends on a large number of complex factors. Rather than employing a structured, tree-based approach to assessing this risk, SAFER characterizes the probability of an explosion for a given PES by assigning the PES to one of twelve logarithmically spaced probability bins. The “least likely” bin represents a one-in-one million probability, while the “most likely” represents a three-in-ten probability. Each bin represents a square root of 10 (approximately 3.16) multiplicative increase in the probability of an explosion.

The assignment of a PES to a particular bin depends on three classes of factors. The first is the primary use of the PES (e.g., training, manufacturing, storage). The second is the type of explosives present, classified into three categories depending on the “compatibility group” (DDESB, 1999) of the explosives. The compatibility groups are based on consideration of explosive characteristics and accident potential. Finally, the probability of explosion assignment is allowed to be altered by a number of so-called scaling factors intended to represent specific characteristics of the PES not addressed by the two-way classification (usage by explosive type). These scaling factors can result in a PES being moved up one or two probability bins, depending on whether it meets any of the scaling factor criteria.

The calculation of the probability of a human fatality given that an explosion occurs is based on an extremely detailed consideration of the different mechanisms of fatality possible during an explosion. The calculation accounts for both the location of the individual at the time of the explosion (inside the PES, between the PES and the ES, at the ES, and inside the ES) as well as the mechanism of fatality (pressure/impulse, glass and building collapse, collateral debris contact, and temperature). In addition, it accounts for the net explosive weight present at the PES as well as details about the construction of the PES and the ES (e.g., percentage of glass and type of roof material). The degree of specificity and detailed consideration of factors used to calculate the probability of a fatality is sufficiently great that a comprehensive treatment is not possible here; we provide only a brief summary.

The fundamental structure of the calculation is based on assuming that four different mechanisms of death—pressure/impulse, glass and building collapse, collateral debris contact, and temperature—act independently to determine whether an individual dies. The constituent probabilities of dying from each of these four causes are the result of complex calculations based on assumptions about the relevant physical processes operating during the explosion. These probabilities are then combined via a single formula to obtain the probability of fatality.

The annual expected exposed population is estimated by first classifying individuals into three personnel categories: those whose jobs relate to the PES (related), persons who are exposed by virtue of em-

ployment (nonrelated), and all others not included in the previous definitions (public). A value for the annual expected exposed population is calculated for each ES in the analysis, based on the number of people expected to be at the ES and their expected durations of stay, with exposure “calculated by multiplying the number of people by the percentage of time they are at the site during the year” (DDESB, 2000). The calculation is performed separately for each of the three groups of individuals. This is because the tolerable level of risk may be higher for those who voluntarily work at the PES with full knowledge of the associated risk compared to that for the public at large.

Although SAFER was designed to assess risks in a framework fundamentally distinct from the UXO problem, selected features may be worth considering in the development of a UXO explosion risk assessment method—especially those features related to characterizing the nature of the explosives. For example, the differentiation by volatility of different kinds of explosives (similar to the “hazard factor” used in OECert) may be relevant to UXO.

ASSESSING MULTIPLE RISK END POINTS

As Chapter Three explained, one of the difficulties encountered in previous attempts to design UXO risk assessment methods is the different characteristics of the risks due to explosion potential of UXO and those due to munitions constituents. Previous attempts to integrate these two different hazard types in a single UXO risk assessment paradigm have been unsuccessful. The ERPS method described above illustrates a mechanism for integrating different risk end points in a single score for the purpose of prioritizing sites. The Army Chemical Stockpile Disposal Program (CSDP) risk assessment for the Tooele Chemical Agent Disposal Facility provides a useful example for integration of multiple risk end points in a site-specific risk assessment.

Chemical weapons are stockpiled and are being (or will be) destroyed at eight sites in the continental United States, including Tooele. These weapons contain either neurotoxic (nerve) agents or mustard (blister) agents. Neurotoxic agents cause deadly changes to the nervous system upon acute exposure. Mustard agents cause severe blistering of the skin and are associated with increased cancer

risk at low concentrations; at high concentrations, they are acutely lethal. The chemicals are stored in a variety of containers, including bulk vessels, rockets, projectiles, mines, bombs, cartridges, and spray tanks.

The chemical weapons disposal process has both acute and chronic risks. Acute risks arise from the potential for handling mistakes, machinery malfunctions, natural events such as earthquakes, and other scenarios to cause an accidental release of neurotoxic or blister agents. Chronic risks result from stack emissions from the incinerators in which the weapons are destroyed.

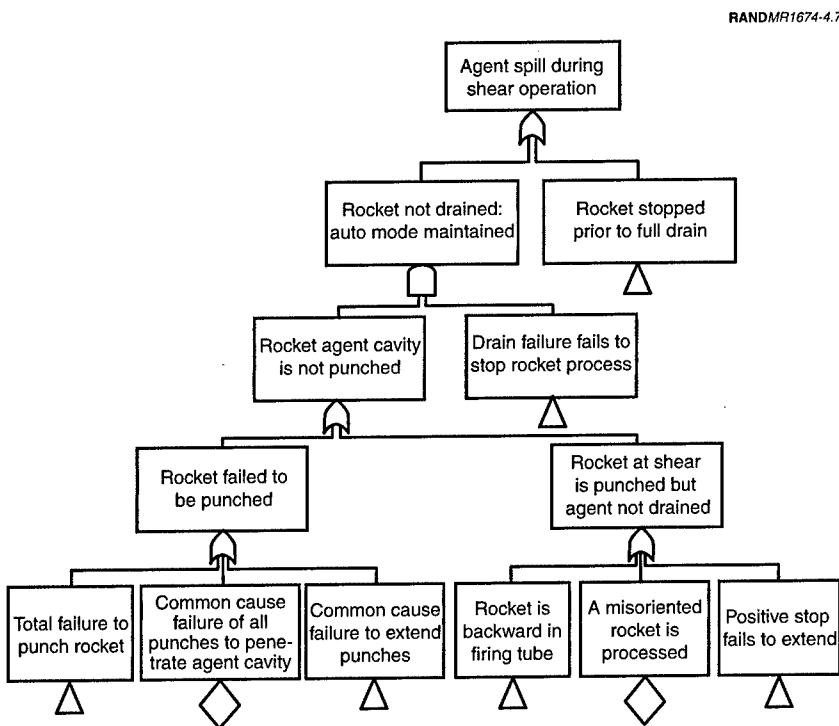
The CSDP used a two-part approach to assessing risks at the Tooele location: (1) EPA methods for assessing chronic risks from airborne contaminants in smokestack emissions, and (2) PRA to estimate acute risks from potential scenarios that could lead to chemical agent releases. The outputs of both methods were such that the risk levels could be compared: all outputs were expressed in terms of probabilities of adverse effects.

The computation of latent health and environmental risks from operation of the incinerators followed prescribed EPA protocols for obtaining operating permits for incinerators under the Resource Conservation and Recovery Act. Thus, the Army did not invest in developing a new method for assessing the effects of smokestack emissions but instead relied on EPA's long history of experience in assessing these types of risks. The Utah Division of Solid and Hazardous Waste, which implements the Resource Conservation and Recovery Act program in Utah, carried out the mechanics of the assessment, which lent additional credibility to the results. The output provided worst-case estimates for increased cancer risk to an adult resident, child resident, subsistence fisher, and subsistence farmer.

To determine acute risk, CSDP developed a site-specific PRA, called the Quantitative Risk Assessment (QRA). The framework for the QRA is based on a logical description of the disposal facility's operation and the development of detailed scenarios that could lead to accidental releases. The QRA first carefully diagrams the entire disposal process. Then, for each step in the process, it identifies possible deviations from normal operations. The factors that cause these deviations are then mapped out in fault trees. Figure 4.7 shows an

example of one of the many fault trees developed for the QRA. In this case, the off-normal scenario, indicated at the top of the tree, is a spill of a chemical agent during one step (known as shear operation) in the process of dismantling rockets containing chemical weapons.

The QRA produced three kinds of output: (1) the probability that a given number of deaths will occur during the seven-year period of planned operation for the disposal facility, (2) the average number of deaths expected over this period, and (3) risk as a function of distance from the site. It provided separate risk estimates for workers

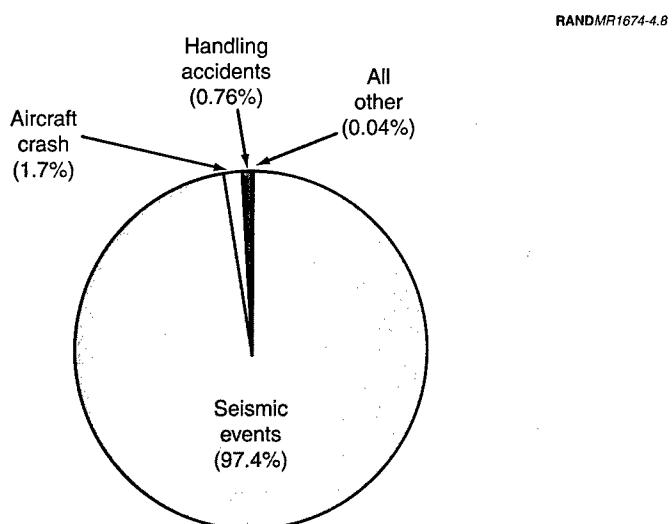


SOURCE: National Research Council (1997b).

Figure 4.7—Example of a Fault Tree Used in Assessing Risks at the Tooele Chemical Agent Disposal Facility

and members of the general public. In addition, the CSDP used the QRA to identify the most significant underlying sources of risks, in order to plan appropriate mitigation strategies. Figure 4.8 shows the contributors to the average public fatality risk, as determined from the QRA. A similar process could be developed to determine the most significant contributors to human health risks at UXO sites.

The CSDP process for Tooele illustrates how EPA methods for assessing chronic risks from dispersed contaminants in the environment can be integrated with PRA methods for assessing acute failure events in a single risk assessment. This approach could serve as a useful paradigm for site-specific assessment of UXO risks: RAGS could be used to assess risks from munitions constituents, and a separate, PRA-based method could be used to assess explosion risks.



SOURCE: National Research Council (1997b).

Figure 4.8—Contributors to the Average Public Fatality Risk Due To Chemical Weapons Storage and Disposal at the Tooele Chemical Agent Disposal Facility

EVALUATING UNCERTAINTY IN RISK ESTIMATES

Uncertainty analysis is fundamental to risk assessment. Risks can almost never be determined with such precision that a single number is adequate to represent all the risk possibilities. None of the UXO risk assessment methods reviewed in Chapter Three provides a mechanism for quantifying uncertainty. However, uncertainty estimates are essential for sound decisionmaking. It is possible that the mean risk for a site could appear moderate, for example, but that the uncertainty is so large that there is significant potential for catastrophic events to occur. Decisionmakers need to know the potential for deviations from the “average” risk.

Uncertainty in risk calculations arises from many sources (Morgan and Henrion, 1990; National Research Council, 1996; Bernstein, 1996). Morgan and Henrion (1990) outline a framework for classifying the sources of uncertainty and approaches for handling it in different scenarios. A few of the major classes of uncertainty include statistical variation, systematic error, variability, randomness, and disagreement. Statistical variation is uncertainty that results from random measurement errors of a quantity. Systematic error is the result of biases in measurements or theories, which, unlike statistical variation, cannot be reduced through repeated observations. Variability is the natural fluctuation of a variable over time, space, or across a population. For example, variability exists in the average height of students in a first grade class. Randomness is uncertainty that is irreducible even in principle. It can be distinguished from variability in that it does not merely result from sampling of a frequency distribution. Whereas further specifying the population, time, or space being considered could reduce variability, it would not reduce randomness. Some argue that randomness is indeterminacy resulting from a limited understanding of the world. Thus, in principle one person might consider a process to be random and another consider it to be deterministic. Finally, disagreement among experts also leads to uncertainty. The disagreement may be on the magnitude of a model parameter, the sign of the model parameter, or even the structure of the model itself.

Uncertainty can be addressed using parametric analysis, descriptive statistics (e.g., confidence intervals), or statistical simulations. The chosen approach depends on both the nature of the problem and the

use of the analysis. Examples of each approach exist in the risk assessment methods reviewed for this chapter.

The SAFER risk assessment approach explicitly examines uncertainty in one of the key model inputs, namely, the amount of explosives present at the PES. To this end, it augments the fundamental risk indicators with two additional metrics intended to provide risk managers with some “worst-case” scenarios to aid in making siting decisions. The first is the maximum possible expected fatalities and the maximum individual probability of fatality, both of which are obtained by assuming the maximum possible amount of explosives. This form of parametric analysis improves decisionmakers’ understanding of the range of possible outcomes when an important parameter cannot be specified precisely. SAFER exemplifies how “bounding analysis” (Morgan, 2001) helps to identify the extremes of potential outcomes when information is too limited to provide a basis for probabilistic analysis.

When sufficient data exist about the sources and effects of hazards, descriptive statistics can be used to characterize uncertainty about the risks. For example, OSHA standards for carcinogens estimate a potency factor for a chemical based on extrapolations of laboratory animal studies to effects in humans. The potency factor translates a specified chemical exposure and dose into an expected response (i.e., risk of cancer). Because there is always uncertainty in the laboratory studies (e.g., due to extrapolating from high to low doses, from animals to humans, or both), the review process for OSHA standards considers the best estimate (mean) of the potency factor and the variance in this potency factor. Statistical confidence intervals are a useful way of reporting variance in the estimated parameters. In cases where uncertainty stems from expert disagreement, not variability or statistical error, descriptive statistics can also be used to characterize subjective judgments about model parameters or risk outcomes (Morgan and Henrion, 1990).

If uncertainty in system components can be characterized using probability distributions, statistical simulations can propagate component uncertainty into uncertainty about the complete systems. For example, the DOE’s TSPA for Yucca Mountain, Nevada, expresses individual potential risk in terms of expected (mean) millirems of exposure to radioactivity per year. The TSPA method handles uncer-

tainty in the parameters of underlying process models by assuming probability distributions for some of the key parameters and then sampling from those distributions, using Monte Carlo methods, to generate hundreds of realizations of TSPA dose rates for each of the scenarios. This provides decisionmakers with mean estimates of exposures and with the expected variance in these estimates (e.g., the 5 and 95 percent confidence limits around the mean). Similar approaches to quantitative evaluation of uncertainty have been developed for PRA methods (Haimes, 1998). One limitation of the statistical approach to uncertainty analysis is that it does not account for uncertainty in the underlying system model.

STANDARDIZING THE RISK ASSESSMENT PROCESS

One consistent characteristic of successful applications of risk-based decisionmaking is standardization. As risk assessment becomes more widely used in an organization's decision process, standardization helps ensure consistency with respect to decisionmaking across people, geographic areas, and time. The benefits of standardization do not stem from adoption of a one-size-fits-all analytical approach. Rather, standardization helps bring consistency and repeatability to risk-based decisionmaking (National Research Council, 1983). Standardization can also increase the transparency of the risk assessment. Formal guidelines for the practice and application of risk assessment help analysts, reviewers, and stakeholders better understand the analysis process. Standardization is important for UXO sites because of the large number of sites involved and the large number of jurisdictions in which those sites are located.

HRS provides one example of how an agency provides for standardization in risk assessments across a program. To promote consistency and comparability of HRS scores across different sites, EPA provides explicitly detailed instructions for HRS implementation. There are several key sources of these instructions. The original reference is the *Federal Register* notice (known as the National Contingency Plan) that describes procedures for implementing CERCLA (EPA, 1982). The more recent *Hazard Ranking System Guidance Manual* provides a more user-friendly presentation of the scoring system (EPA, 1992). This manual includes hypothetical examples to clarify nuances in the scoring process. In addition to written guide-

lines, the EPA has developed a number of electronic resources. One of these is an online version of the HRS training course, available on the EPA web site. Another is a software set, also available on the web. For example, the software package "Preliminary Assessment Score" guides the user through a considerably simplified version of the HRS based on some worst-case scenario assumptions and a relatively small amount of data that could be available from a preliminary assessment of a CERCLA site.

Within DoD, SAFER provides an example of a standardized risk assessment process. Software for performing the SAFER calculations is available online (DDESB, 2002). The documentation provides complete instructions on how to calculate the necessary input quantities, which are flexible enough to be applied to different site conditions. Given that SAFER is intended to produce credible estimates of expected fatalities, it relies on a number of highly technical calculations. Most of these technicalities are transparent to the user. Model implementation with the software should be reasonably straightforward for site managers. Although SAFER requires a number of inputs, the vast majority of these inputs are not subject to disagreement among raters. Thus, SAFER output should be relatively stable among different assessors who use the same information about a site.

For UXO sites, the adoption of uniform risk assessment standards with specific instructions will help promote consistency in results and transparency to stakeholders.

WORKING WITH STAKEHOLDERS

With the realization that value-based decisions are often inextricable from the risk characterization and assessment processes (National Research Council, 1996), risk managers have worked toward methods of incorporating stakeholder participation into decisionmaking. In fact, recent research provides evidence of stakeholder participation leading to higher-quality decisions, citing increased joint gains, lower-cost outcomes, and incorporation of innovative ideas (Beierle, 2002). Given the multiobjective nature of UXO risk assessment and the lack of consensus on what represents acceptable cleanup levels, stakeholder participation will be critical to successful UXO risk management.

The failure in implementation of the technically strong DOE ERPS model, described above, is an illustration of what can go wrong when public participation is not addressed adequately. At the outset of ERPS development, the DOE appointed a group of stakeholders, the “External Review Group” (ERG), to participate in system development. The membership of this group included representatives from states with DOE installations, Native American tribes, the National Governors Association, the National Conference of State Legislators, the National Association of Attorneys General, the EPA, the Natural Resources Defense Council, and the Environmental Defense Fund. The ERPS design team surveyed ERG members about their key priorities for the system and held two initial meetings with the group.

From the beginning, the ERG meetings were contentious. This was in large part because some ERG members were concerned that DOE was going to use ERPS to back out of negotiated agreements with states and EPA that specified certain cleanup requirements for the contaminated sites. Nonetheless, the design team worked to incorporate ERG concerns in the ERPS design. For example, they developed a screening system to ensure that sites posing an imminent risk would be guaranteed funding; essentially, such sites were eliminated from prioritization and placed at the top of the list. Also, they developed a constraint that would allow consideration only of funding allocations that would ensure full regulatory compliance. But before ERG had had a chance to review the final ERPS design and before all of ERG’s concerns were fully addressed, DOE made an executive decision to terminate ERG. In a retrospective analysis of the failure of ERPS implementation, Jenni and others involved in the ERPS design described DOE’s decision:

Citing the various changes that had been made in response to ERG comments, the DOE declared that the ERG had completed its task and no future meetings would be scheduled. Public involvement in future applications of ERPS was promised; however, responsibility and the specifics for public involvement were left to local field office and facility personnel (Jenni et al., 1995, p. 405).

Jenni et al. attributed the failure of ERPS in part to the lack of effective DOE follow-through on the promise for public involvement. Jenni et al. concluded, “DOE asked for the opinions of the ERG but did not respond to them, . . . thereby permanently alienating the

ERG.” Further, they concluded, “commitment to public involvement was expressed and verbal promises to publish a public participation plan were made, but there was never adequate follow-through.”

In contrast to DOE’s practices during the development of ERPS, community involvement is formally required as part of the EPA RAGS process. EPA’s original guidelines for community involvement in risk assessment are specified in RAGS Part A, Chapter 9. Later, EPA released a supplement to RAGS Part A to augment the information in Chapter 9. The supplement is fundamentally about how EPA can improve relations with the public, because many communities affected by past CERCLA remediation efforts were unsatisfied with EPA’s approach. The RAGS documentation, along with DoD’s own internal policies for community involvement at sites in the Defense Environmental Restoration Program, could provide useful examples for the development of guidelines for community involvement in UXO risk assessment.

One of the important features of effective community involvement plans (in addition to following through on them, which DOE failed to do with ERPS) is to give the community a role in designing the risk assessment and data-collection protocol from the beginning, rather than after the risk assessment is completed. Substantial research has documented that community involvement is most effective when it begins very early and that failure to do so can lead to major road-blocks later in the process (National Research Council, 1996, 1997a). For example, while EPA guidance stresses the importance of public participation at CERCLA sites, in the past some community members complained that they were not allowed to be involved until after a set of alternative cleanup options had been identified (National Research Council, 1997a). In a review of barriers to implementing innovative technologies in the cleanup of contaminated groundwater, the National Research Council suggested that the lack of sufficiently early community involvement has contributed to slowing the remediation process and has limited the potential for selection of innovative and potentially cost-saving approaches (National Research Council, 1997a). The tree-based methods of PRA could provide effective vehicles for early community involvement: community groups could be asked to help identify scenarios of concern that could be diagrammed and analyzed with fault and/or event trees.

Experience from DOE and EPA illustrates that involving the community early in the development of a risk assessment process and that following through on the commitment to that involvement are critical ingredients of successful risk assessment.

DEVELOPING TRUST AMONG STAKEHOLDERS

Even simple risk assessments can appear complicated to the common observer. Each of the analytical approaches presented in this chapter relies on collecting large amounts of data. They all also incorporate many types of subjective judgments, sometimes including the relative importance of different risk attributes and selection of the risks considered. Many consider risk assessment to be a science, in part because analysis usually incorporates observed data. However, it is important that risk assessment results are not presented as facts developed from an objective model (National Research Council, 1996). In fact, masking the subjective judgments inherent in a risk assessment can undermine the utility of the results in a public policy context.

The best way to counter this concern is to incorporate transparency and independent reviews into the risk assessment process. For a risk assessment to be transparent, all underlying assumptions must be clearly stated. This includes not only the sources and rationales about selection of parameter values but also decisions about model structure and the inclusion or exclusion of risks to analyze. A benefit of the tree-based, PRA tools that the NRC, NASA, FAA, CSDP, and others use is that they can provide a transparent framework for illustrating the risk analysis process. For example, event and fault trees can illustrate the particular risk scenarios considered and the combinations of events the analyst believes are necessary for those scenarios to unfold, along with the probability of each scenario and underlying event. Thus, stakeholders reviewing the risk assessment will have a clear picture of the analyst's thought process. Stakeholders could also be involved in identifying the scenarios considered in the trees and the connections among events leading to the scenarios.

Independent review reinforces the requirements for transparency because review panel members must be able to make an adequate evaluation of the risk assessment methods used. Depending on how the review panel is assembled, external review can lend further

credibility to the assessment process and potentially build trust in the results with stakeholder groups. For example, the CSDP relied on multiple, rigorous external reviews in developing the Tooele risk assessment. The Army formed an independent expert panel, the “Risk Assessment Expert Panel on the Tooele Chemical Agent Disposal Facility,” to carry out detailed, step-by-step reviews of the QRA as it was being developed by an Army contractor. The five-member panel included nationally recognized experts in risk assessment from academia and industry. In addition, the National Research Council reviewed the method once it was completed. This expert review was intended to improve the technical soundness of the risk assessment method and increase the likelihood that its application and output would be widely accepted by stakeholders.

UXO risk assessment will not be simple. Multiple end points and great uncertainty complicate the task. Transparency and independent review can help decisionmakers earn public trust and support for risk management decisions.

CONCLUSIONS

Risk assessment is widely practiced in the federal government, not just for environmental regulation but also for managing the safety of airlines, space shuttles, nuclear power plants, and workplaces, as well as for many other purposes. Risk assessment tools used in several other agencies have been developed and refined over decades. These existing tools provide examples that can guide the development of improved UXO risk assessment methods. Table 4.5 summarizes the common problems in risk assessment that the Army and DoD are now encountering at UXO sites and identifies the methods from among those we reviewed that provide the best example solutions.

Table 4.5

Summary: Existing Federal Methods That Illustrate Solutions to Common Problems in Risk Assessment

Risk Assessment Problem	Example Solutions
Prioritizing sites for remediation	HRS RRSE
Assessing (quantitatively) site-specific risks	RAGS PRA
Assessing multiple-risk end points	CSDP Tooele risk assessment
Evaluating uncertainty	TSPA SAFER OSHA carcinogen risk assessments
Standardizing the process	HRS SAFER
Working with stakeholders	RAGS
Developing trust among stakeholders	CSDP Tooele risk assessment

Chapter Five

CONCLUSIONS AND RECOMMENDATIONS

None of the UXO risk assessment methods that we evaluated fully meets the Army's need for sound technical analysis to inform decisionmaking, either for the purpose of setting priorities among UXO sites or for detailed analysis of explosion and munitions constituents risks at individual sites. Table 5.1 summarizes the methods' strengths and limitations. As shown, each method falls short in one or more of the key criteria necessary for an effective method: technical soundness of risk calculations, effectiveness of implementation, or ease of communication. Furthermore, stakeholders and regulators involved at the various sites have not uniformly accepted these methods as credible elements of the decisionmaking process, and continued reliance on them is likely to delay the UXO response process further.

A fundamental reason why none of the modeling methods evaluated meets the Army's needs is that the UXO problem is not reducible to a single, objective measure of risk. Risk methods must address the risk of explosion of the munitions but also consider the risk of chemicals from exploded munitions and UXO that leach into the soil and groundwater. Further, the methods used for analyzing these two broad categories of risk (explosion and munitions constituents), while different in substance, both depend on subjective judgments about modeling assumptions and data. For example, assessing the explosion risk requires, among other types of information, estimates of the probability that humans will come in contact with UXO. These estimates require assumptions about human behavior and predictions of future population and land use; the density and distribution of UXO items that cannot be seen because they are buried; and the

Table 5.1
Summary of UXO Risk Assessment Methods

Method	Purpose	Pros	Cons	Summary Evaluation
IR3M	Assess explosion and munitions constituents risks at individual sites	Simple output	Output does not always correlate to risk Output can mask important risk information Decision rules not technically justified Basis for input values not justified Not always reproducible Does not address uncertainty Data requirements insufficient to reflect problem complexity Instructions unclear	Multiple technical weaknesses; should not be developed further
OECert	Developed to prioritize risks among sites, but in practice used to assess explosion risks at individual sites	Comprehensive modeling of exposure process Analytical process used to determine explosion potential of different munitions Adaptable	Does not address munitions constituents risk Exposure models not validated Many exposure assumptions not justified Uncertainties not addressed Calculations not presented clearly Not easily communicated to stakeholders Lack of stakeholder involvement in developing exposure assumptions	Elements of the method (exposure models, UXO categorization method) might form part of future risk assessment method but would need much refinement

Table 5.1—continued

Method	Purpose	Pros	Cons	Summary Evaluation
RAC	Prioritize among sites	Appears logically sound Assumptions clearly explained Reproducible Practical (data requirements suitable for purpose) Adaptable	Does not consider munitions constituents Does not address uncertainty Basis for some assumptions not provided Does not address munitions constituents risk Risk model relation to actual magnitude of site risk unknown Assumptions not explained Uncertainty not addressed Not reproducible Data requirements too minimal Results easily manipulated	Well suited for purpose, but only addresses prioritization of explosion risk Has many limitations and should be discontinued
OERIA	Assess explosion risk at individual sites	Easy to use Adaptable	Focuses exclusively on ecological risks Does not consider munitions constituents Assumptions not justified Uncertainties not addressed Instructions somewhat unclear	Meets need to identify UXO sites whose cleanup is complicated by regulatory requirements related to natural or cultural resources, but needs substantial further development and validation
NCRB	Prioritize ecological risks among sites	Appears to be reproducible Adaptable		

probability that the fuze of a UXO item that may have been buried for decades is intact.

Despite these difficulties, credible UXO risk assessment methods are needed to allow progress toward defining acceptable UXO cleanup standards. Most would agree that zero risk at UXO sites would be the ideal standard, but in reality that standard cannot be achieved with the resources and technical capabilities available now or in the foreseeable future. As explained in Chapter One, the only process that currently can guarantee that all UXO has been removed involves

- burning or cutting all vegetation,
- excavating the entire site one foot at a time down to the maximum possible penetration depth of the UXO (as much as 10 feet or more), and
- sifting all the excavated soil.

This process is too costly to be feasible for the DoD to implement as the standard process for UXO response. Furthermore, it causes irreparable damage to what are often uniquely preserved ecosystems and in many cases will be unacceptable to regulators from natural resource management agencies.

Because sifting is neither possible nor desirable in most cases, UXO clearance relies on metal detectors to locate buried UXO items. Metal detection technologies are imperfect: they do not find all buried UXO. The Army and regulators alike need to publicly acknowledge this reality and design risk-informed decisionmaking processes that can lead to acceptable compromises.

This chapter suggests steps for the Army and DoD to take toward developing risk assessment methods for prioritizing UXO sites and for evaluating individual sites in detail. The goal is to develop methods that are technically credible, acceptable to stakeholders, and practical to implement. While challenging, evidence from the successful use of risk assessment in other agencies (described in Chapter Four) demonstrates that the task is not impossible.

RECOMMENDED PROCESS FOR PRIORITIZATION

We concluded in Chapter Three that a prioritization method for UXO sites should not have as its output a single metric that combines explosion risks and munitions constituents risks. This is because any attempt to integrate acute and chronic risk into a single risk value requires value judgments about the relative importance of the different risks. While such value judgments are part of an overall decision-making process, they should enter only at explicit decision points and not be masked in the technical analysis of risk (National Research Council, 1983, 1996). Therefore, we recommend a two-stage process for prioritizing UXO sites, with information about explosion risk magnitude and munitions constituents risk magnitude preserved in the final output.

Figures 5.1 and 5.2 illustrate the recommended process. In the first stage, sites are sorted into “bins” or classes based on explosion risk. Within each explosion risk bin, sites are sorted according to risks from munitions constituents. We suggest using explosion risk as the first filter for sorting sites because it presents the most immediate hazard. Failure to clear explosive items could preclude other UXO response actions, such as installation of systems to monitor and treat groundwater and soil contamination. In addition, explosive items are sources of potential munitions constituents contamination of soil and groundwater. Even if all contaminated soil and water were cleaned, a risk of further contamination would remain as long as UXO is present.

Process Description

As shown in Figure 5.1, the prioritization process begins with an evaluation of existing information (historical records, interviews, site surveys, reports of encounters with UXO, and so on—information typically collected for what is known as an “archives search report”) to determine whether explosion risks are present. If such risks exist, the second step involves evaluating whether available data are sufficient to proceed with estimating the risks associated with potential exposure scenarios. The shaded box to the left of the second tier indicates that before the prioritization process can proceed, technical criteria for data sufficiency must be established and satisfied.

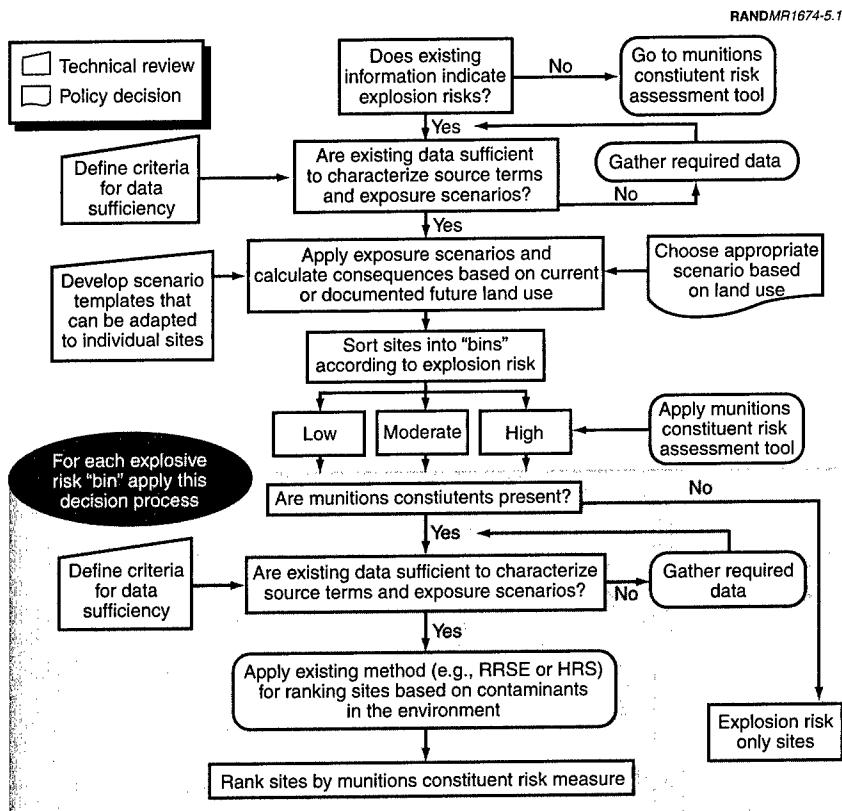


Figure 5.1—Recommended Two-Stage Prioritization Process

The next critical step in the process, shown in the middle section of Figure 5.1, involves sorting sites into “bins” according to the level of explosion risk. The figure shows three bins, labeled “low,” “moderate,” and “high” consequence. However, any number of bins could be used, and the labels could be numerical (1, 2, 3, 4, and so on) or have other ordinal representations. The key is that sites are grouped according to level of explosion risk, with this risk being estimated by a method that calculates the probability and consequences of explosion based on considering possible ways in which people might be exposed to UXO. RAC, which prioritizes sites based on explosion risk on a scale of 1 to 5, could provide a starting point

for this method, but some modification (for example, assumptions about exposure) may be needed to reflect stakeholder concerns. The process that SAFER uses to sort sites into 10 bins based on explosion probability and then to estimate exposure potential for nearby populations might provide a useful example of a sorting method. Whatever process is chosen should not be overly detailed and should rely as much as possible on readily available information so that it can be implemented efficiently at the programmatic level. Nonetheless, it should satisfy the evaluation criteria for an effective risk assessment process that are described in detail in Chapter Two and summarized in Table 5.2.

The next step of the process accounts for risks from munitions constituents in the environment and is illustrated in the bottom of Figure 5.1. As shown, sites within each explosion risk bin are sorted ac-

Table 5.2
Summary of Criteria for Evaluating UXO Risk Assessment Methods

Category	Criteria
Risk calculation features	Comprehensive Logically sound Consistent with current science Models well defined, validated Assumptions documented, reasonable, used consistently, eliminated when unnecessary Uncertainties described accurately Output reproducible Data sources and calculations clear Peer reviewed
Implementation features	Practical Technical instructions clear Broadly applicable (i.e., to diverse sites) Free of loopholes
Communication features	Stakeholder input points clear Transparent Uncertainty clearly communicated Conducive to education

cording to munitions constituents risk. Two existing methods for such sorting are widely used, and we recommend that one of them be selected for this purpose. The two methods are the Defense Environmental Restoration Program's RRSE framework and the EPA's HRS. The relevant features of these methods are described in Chapter Four. Either of the two could help to rank UXO sites in the explosion risk bins according to risks from munitions constituents.

Figure 5.2 summarizes the output of the ranking process. The process would result in bins of sites with similar degrees of explosion risk. In Figure 5.2, we show four bins (no, low, moderate, and high explosion risk), but any reasonably small number could be used. Within these bins, sites are ranked again according to risks from munitions constituents.

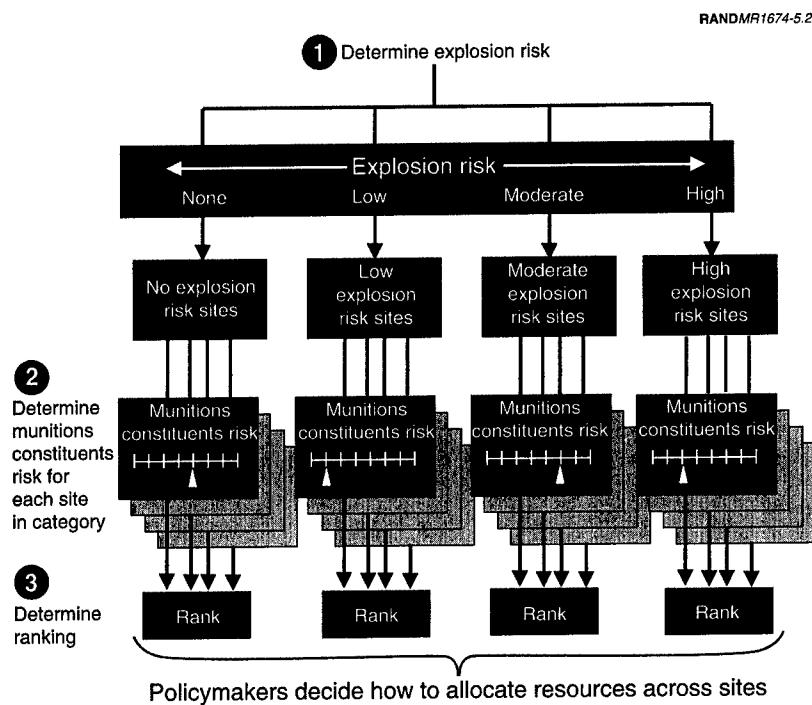


Figure 5.2—Summary of Recommended Prioritization Process

As discussed in Chapter Three, there is no technical basis for assigning to a site a single risk measure that combines explosion and other constituent risks. As a result, a policy decision needs to be made about how to allocate resources among sites with high explosion risk but low risk from munitions constituents versus sites with low or no explosion risk but high risk from munitions constituents.

Oversight by Technical Review Board and Advisory Committee

We recommend that the Army appoint an independent technical review board to guide technical development and implementation of the prioritization process. The review board should consist of non-partisan technical experts in exposure and explosion risk modeling. The technical board would provide a critical technical review of the prioritization process and its appropriate application. In addition, it would establish requirements for input data. If desired, review panels could also be constituted for each site to review the prioritization of each site once the prioritization framework is in place. A site-specific panel could be more accessible to the interested public and work closely with site managers on their particular technical questions.

Site prioritization needs to be carried out to the extent possible with readily available information. The detailed surveys required for site-specific risk assessment usually will not be available at the prioritization stage. Without detailed data, quantifying uncertainty is difficult. The technical review board could assess the quality of input data for individual sites and the associated level of uncertainty in the output of the prioritization in order to determine whether additional data need to be gathered to increase confidence in the results. For example, the review board could take on the question of how to judge the validity of assumptions about UXO density at a given site.

In addition to the technical review board, we recommend that the Army establish an advisory committee consisting of citizens who live near or at UXO sites, state and federal regulators, Army officials, and technical experts. The advisory committee would work closely with the Army from the start of the development process to consider appropriate choices about land use scenarios, essential features of risk

communication, and other implementation features associated with the prioritization framework. Membership should not overlap with the technical review board.

A number of policy decisions need to be made in conjunction with the development of a prioritization process. Figures 5.1 and 5.2 highlight some of the key decisions, but there are others as well. Box 5.1 provides examples.

Two Lists, Based on Land Use

The risk of people suffering injury from UXO depends on how the land will be used. Land use information is therefore critical for the prioritization process. However, for many sites, land use has not been decided formally. Where future land use is undetermined, there is a risk that some decisionmakers might manipulate the prioritization process to favor their site by choosing, for example, the most risky land use (such as child care)—even if that use is highly unlikely or ill advised.

To reduce the likelihood of manipulation of the prioritization process, we recommend maintaining two priority lists, both developed using the framework shown in Figures 5.1 and 5.2. The first list would include all sites and would be based only on current land use.

Box 5.1. Policy Decisions To Be Considered by Recommended UXO Advisory Committee

What exposure scenarios should be considered for different land uses?

Should ecological risk be considered more fully in the prioritization stage? If so, how can these risks be compared across sites?

What types of data should trigger reconsideration of a site's rank?

What is the relative importance of high explosion risk versus high munitions constituents risk?

How should uncertainty be handled in the prioritization process? For example, should the riskiest possible scenario or the most likely possible scenario guide the assessment?

The second list would include only sites for which future land use has been decided. Inclusion on the second list would require legal documentation of the future use, such as a locally approved land use plan for the area that has already received DoD concurrence in the case of BRAC sites. For FUDS, a policy decision would need to be made as to whether future land use should be constrained by current land use. Policymakers could decide how much funding to allocate to sites that pose an imminent risk due to current use and how much to allocate to sites that pose a significant risk in the future, based on proposed and officially documented new land uses. This approach would reduce the susceptibility of the prioritization process to manipulation.

RECOMMENDED SITE-SPECIFIC RISK ASSESSMENT PROCESS

Unlike risk estimates for the purposes of prioritization, which must be comparable between sites, site-specific risk assessment must be tailored to local concerns and conditions. Thus, an omnibus tool applicable to all sites is not advised because such a tool will almost certainly lack the flexibility needed for stakeholder support. Thus, we recommend a site-specific risk assessment process that can be adapted with stakeholder input at each site, rather than a “one-size-fits-all” tool applied uniformly across all sites. The process would need to be supported by technical methods that can be adapted to different sites, much as the CERCLA risk assessment process is supported by technical methods for exposure assessment and dose-response assessment. As we described in Chapter Three, the available methods for assessment of explosion risks do not satisfy the technical criteria required of a risk assessment, and thus a new method must be developed.

The probabilistic approaches used by the NRC, CSDP, NASA, and the FAA provide possible models for the assessment of UXO explosion risks. All of these agencies employ the same basic approach (see Chapter Four). The strategy involves defining sources of risk, mapping the scenarios that could lead to adverse consequences, and then using probabilistic methods to determine the likelihood of consequences occurring. A similar strategy could be applied to assessing explosion risks at UXO sites.

As is the case for site prioritization, available methods are suitable for characterizing munitions constituents risks at the individual site level. The problem of characterizing low concentrations of munitions constituents in the environment is no different from that of characterizing distributions of nonexplosive contaminants found at nonmilitary waste sites. RAGS, described in Chapter Four, has long been used and is widely accepted for this purpose. Consequently, the Army need not create a new method for munitions constituents. In fact, an attempt to do so would most likely be questioned because of the existence and regulatory acceptance of RAGS.

Description of Recommended Approach

The PRA approach used by the NRC, CSDP, NASA, and FAA begins by asking the questions "What can happen?" and "How can it happen?" As explained in Chapter Four, the answers to these questions form the basis for event and fault trees, as well as for associated probability calculations. Each event tree illustrates one potential undesired operational scenario and the chain of adverse events that would ensue. A fault tree illustrates an undesired end state and the chain of events needed to trigger it.

In developing OECert, Army engineers attempted detailed modeling of the amount of area that persons engaged in particular activities on UXO sites would cover (see Chapter Three). Some of the information from this effort may be useful in developing fault or event trees for UXO risk assessment. However, the scenarios, assumptions, and approximations would need to be modified substantially with stakeholder and technical input.

The PRA approach has a number of advantages. It requires analysts at the site to decompose the elements of risk systematically and to construct a formal structure for analyzing the risk. This process allows for the identification of "risk drivers" (dominant sources of risk) and the planning of optimal strategies for mitigating them. The logical structure and graphical presentation facilitate stakeholder input and communication. These features can help shift the debate from circular arguments over what level of remaining UXO is acceptable to specific questions such as the likelihood that a UXO fuze will be intact, the probability that a child will enter a former UXO area (based on past experience), and the probability of detecting UXO (given

performance information for the metal detectors used at the site). In addition, the approach is easily adapted to new information about event sequences or probabilities. Finally, the results can be compared directly to risks of adverse health impacts from munitions constituents, because the approach quantifies risk values as explicit probabilities, rather than as dimensionless rankings (as was the case for IR3M and other existing UXO risk assessment methods). Although some stakeholders have objected to any quantification of risk on the grounds that no level of risk above zero is acceptable, as explained in Chapter Two, quantification of risk does not imply that a particular risk level is acceptable. Rather, quantitative risk values are but one input to complex risk decisions. The more technically sound these risk values, the more sound the basis for the decisions.

The PRA approach is not without difficulties. The method is easiest to use for engineered systems, such as a space shuttle or a weapons storage facility, for which past performance information about specific system components is generally more complete than is the case for UXO sites. At UXO sites, a great deal of uncertainty exists about UXO types, condition, and location. In addition, the means by which people might come in contact with UXO cannot be quantified with certainty or fully controlled. Nonetheless, in the engineered systems for which PRA has been used for decades, significant uncertainties about component failure rates and human behavior also exist, and methods have been developed to incorporate uncertainty into the analysis (Haimes, 1998). These same methods for uncertainty analysis could be employed for assessing explosion risk at UXO sites. In some cases, uncertainty analysis may show that the potential range of risks is so large that a point estimate of risk is not useful for decisionmaking. In those cases, the event or fault tree can identify the basic events that are the major contributors to uncertainty, and more information can be gathered about those events before a decision is made.

Development of Event and Fault Trees

Ensuring that the structure of an event or fault tree (i.e., the events and outcomes) is comprehensive can be difficult (see Chapter Four). Stakeholder and public participation in this process will add transparency and help ensure that outcomes and events of concern are

addressed. For the events in the trees corresponding to the likelihood that a UXO item will detonate, extensive experience and data would support specification of probabilities. Explosive ordnance disposal specialists could be consulted to develop models specific to the UXO problem. Information such as type of ordnance, failure modes leading to malfunction upon initial firing, type of fuze, mass of explosive material, and age of the ordnance would provide inputs.

The parts of the tree representing human contact with UXO would be more difficult to diagram because the assumed encounters are not easy to quantify. Rather, stakeholders would need to decide what kinds of exposure scenarios to consider and how detailed these scenarios should be. We recommend defining exposure scenarios at the local level. Once again, significant participation of local stakeholders in outlining these scenarios would increase the acceptance and utility of the site-specific assessment method in a risk management process. In fact, the structure of event and fault tree analysis provides a useful model for organizing stakeholder deliberations over these difficult policy issues. A set of template scenarios and associated event and fault trees could be constructed at the national level to serve as models for local analysts.

EPA's Federal Facilities Restoration and Reuse Office has raised two important issues associated with estimating explosion risk. First, estimates of explosion risk rely on assumptions about the spatial distribution of UXO items. When archival and field information are limited, Army methods such as OECert have assumed that UXO is distributed homogeneously across a site. This assumption is unlikely to be valid in many circumstances when training actions were directed at specific targets. Second, EPA's Federal Facilities Office told the authors that the modeling of human behavior to estimate probabilities of exposure scenarios is unacceptable. This position implies that risk assessors should assume a probability of one that a UXO item will cause injury or death by virtue of its existence, regardless of the likelihood of exposure. This view represents a significant policy decision whose practical and budgetary implications should be fully explored. Human behavior modeling is critical in many other widely accepted applications of risk assessment. For example, in evaluating the risks of contaminants in potable water, assessors typically assume that individuals will each drink two liters of potable water per day.

Oversight by a Technical Review Board

We recommend that an independent panel of technical experts oversee development of the PRA process for UXO sites, just as a technical review board guided the design of the CSDP QRA process. As was the case for the QRA, an expert panel of advisors would help to ensure the technical quality and credibility of the process. To provide additional quality control and credibility, the Army may want to consider having the final product reviewed by a high-level independent technical organization such as the National Research Council, as was done for the CSDP risk assessment method. The method should satisfy the evaluation criteria described in Chapter Two and summarized in Table 5.2.

SUMMARY OF RECOMMENDATIONS

In summary, UXO risk assessment requires two processes. The first process would prioritize UXO sites for remediation (as Congress now requires). The second process would provide for detailed evaluations of appropriate responses to UXO at specific sites. The Army would benefit from new technical methods for both applications because the existing options are unsatisfactory for the Army's current and future needs.

We recommend that a new UXO prioritization process (1) sort sites into bins by explosion risk and (2) within these bins, sort sites by munitions constituents risks.

The suggested prioritization process would preserve the information about the two separate risk types: although sites would be grouped first according to explosion risk, within these groups the sites would be ordered by munitions constituents risk. Policymakers then could decide how much to allocate to sites with varying levels of explosion versus munitions constituents risks.

We recommend developing a new process for sorting sites by explosion risk (stage one of the prioritization process).

RAC could provide elements and a starting point for the new process, but stakeholder concerns would need to be addressed.

We recommend using HRS or RRSE for sorting sites by munitions constituents risks (stage two of the prioritization process).

These methods are well established and well accepted. There is no need for a new approach for munitions constituents, since the behavior of these contaminants and the risks they pose are similar to those of chemical contaminants found at non-UXO hazardous waste sites.

We recommend producing two separate UXO site priority lists: one for sites with known and documented future land use and another for sites with uncertain future land use.

Having two lists would prevent manipulation of the process by choosing the least restrictive land uses. Also, it would allow policy-makers to decide how to trade off current and future risks when allocating funds. The lists could be updated annually or as often as new information became available.

We recommend using RAGS for site-specific assessment of munitions constituents risks.

RAGS is well established for assessing risks of chemicals in water and soil, and there is no need for the Army to develop a new method.

We recommend that the Army develop a new, probabilistic approach using the PRA techniques developed by the NRC, NASA, and others for site-specific assessment of explosion risks.

None of the available UXO explosion risk assessment methods by itself satisfies our technical criteria for an effective risk assessment method, and therefore a new approach is needed. Many other agencies use PRA to assess risks of acute events analogous to UXO explosion.

Finally, we recommend that an independent technical review board and an advisory committee of stakeholders oversee development of both the prioritization system and scenario-based site-specific risk assessment processes.

The technical board would consist of independent experts in risk assessment and explosive ordnance disposal. The advisory committee

Conclusions and Recommendations 141

would include representatives of the different groups of stakeholders (state regulators, Native Americans, federal regulators, members of the public, military personnel) involved at UXO sites.

REFERENCES

Agency for Toxic Substances and Disease Registry (ATSDR), *Toxicological Profile for RDX*, Atlanta, GA: ATSDR, 1995a.

_____, *Toxicological Profile for RDX*, Atlanta, GA: ATSDR, 1995b.

_____, *Toxicological Profile for HMX*, Atlanta, GA: ATSDR, 1997.

_____, *Toxicological Profile for 2, 4- and 2,6-Dinitrotoluene*, Atlanta, GA: ATSDR, 1998.

Army Corps of Engineers, *Ordnance and Explosives Cost Effectiveness Risk Tool (OECert)*, Final Report (Version E), Huntsville, AL: Ordnance and Explosives Mandatory Center for Excellence, 1995.

Army Engineering and Support Center, *Risk Assessment Procedures for Ordnance and Explosives (OE) Sites*, Huntsville, AL: Army Engineering and Support Center, 2000.

Army Environmental Center, *Range Response: Natural and Cultural Resources Issues, Phase II Draft*, Aberdeen Proving Ground, MD: Army Environmental Center Conservation Division, 2001.

Army Science Board, *Results of the Range Rule Study Group*, Aberdeen Proving Ground, MD: Army Environmental Center, 1998.

Barry, Thomas F., *Risk-Informed, Performance-Based Industrial Fire Protection*, Knoxville, TN: Fire Risk Forum, 2002.

Bechtel SAIC Company, LLC, *Total System Performance Assessment: Analyses for Disposal of Commercial and DOE Waste Inventories at Yucca Mountain—Input to Final Environmental Impact Statement*

and Site Suitability Evaluation, Las Vegas, NV: Bechtel SAIC Co., LLC, 2001.

Beierle, Thomas C., "The Quality of Stakeholder-based Decisions," *Risk Analysis*, Vol. 22, No. 4, 2002, pp. 739-749.

Bernstein, Peter L., *Against the Gods: The Remarkable Story of Risk*, New York: John Wiley & Sons, 1996.

Bryan, W. Ray, and Michael B. Shimkin, "Quantitative Analysis of Dose-Response Data Obtained with Three Carcinogenic Hydrocarbons in Strain C3H Male Mice," *Journal of the National Cancer Institute*, Vol. 3, 1943, pp. 503-531.

Burke, Thomas, J. Dyer, W. Edwards, Karen Hulebak, N. Matalas, H. Mesner, Frank Parker, Glenn Paulson, Milton Russell, and R. Zeckhauser, *Report of the Technical Review Group of the Department of Energy's Priority System for Environmental Restoration*, Washington, D.C.: DOE Office of Environmental Restoration and Waste Management, 1991.

Carson, Rachel, *Silent Spring*, Boston, MA: Houghton Mifflin, 2002.

Covello, Vincent T., and Jeryl Mumpower, "Risk Analysis and Risk Management: An Historical Perspective," *Risk Analysis*, Vol. 5, No. 2, 1985, pp. 103-120.

Dawes, Robyn. M., "Psychological Measurement," *Psychological Review*, Vol. 101, No. 2, 1994, pp. 278-281.

Department of Defense, "Closed, Transferred, and Transferring Ranges Containing Military Munitions: Proposed Rule," *Federal Register*, Vol. 62, No. 187, 1997a, pp. 50, 795-850, 843.

_____, *Relative Risk Site Evaluation Primer*, Washington, D.C.: Department of Defense Interservice Relative Risk Working Group, 1997b.

_____, *Report of the Defense Science Board Task Force on Unexploded Ordnance (UXO) Clearance, Active Range Clearance, and Explosive Ordnance Disposal (EOD) Programs*, Washington, D.C.: Government Printing Office, 1998.

_____, *Preliminary Validation Report: Interim Range Rule Risk Methodology, Appendix D*, Washington, D.C.: Department of Defense, 2000.

_____, *Fiscal Year 2001 Defense Environmental Restoration Program Annual Report to Congress*, Washington, D.C.: Department of Defense, 2002.

Department of Defense Explosives Safety Board (DDESB), *Department of Defense Ammunition and Explosives Safety Standards: Under Secretary of Defense for Acquisition and Technology Directive*, DoD 6044.9-STD, Washington, D.C.: Department of Defense, 1999.

_____, *Risk-Based Explosives Safety Analysis*, technical paper no. 14, Washington, D.C.: Department of Defense, 2000.

_____, *Safety Assessment for Explosive Risk (SAFER)*, Washington, D.C.: Department of Defense, 2002.

Department of Energy (DOE), *Preliminary Design Report: DOE Environmental Restoration Priority System*, Washington, D.C.: DOE Office of Environmental Restoration and Waste Management, 1991.

Edwards, Ward, and J. Robert Newman, *Multi-attribute Evaluation: Quantitative Applications in the Social Sciences*, Troy, NY: Sage Publications, 1982.

Environmental Protection Agency (EPA), "Hazard Ranking System," in *National Contingency Plan*, 40 CFR Part 300, 1982, Appendix A.

_____, *Risk Assessment Guidance for Superfund, Volume I: Human Health Evaluation Manual, Part A*, Washington, D.C.: EPA, 1989.

_____, "Hazard Ranking System, Final Rule," *Federal Register*, Vol. 55, No. 241, pp. 51532-51667, December 14, 1990.

_____, *Risk Assessment Guidance for Superfund, Volume I: Human Health Evaluation Manual, Part B*, Washington, D.C.: EPA, 1991a.

_____, *Risk Assessment Guidance for Superfund, Volume I: Human Health Evaluation Manual, Part C*, Washington, D.C.: EPA, 1991b.

_____, *Hazard Ranking System Guidance Manual*, Washington, D.C.: EPA, 1992.

_____, *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments, Interim Final*, Washington, D.C.: EPA, 1997a.

_____, *Representative Sampling Guidance Document, Volume 3: Ecological (Draft)*, Edison, N.J.: EPA, Office of Emergency and Remedial Response, Environmental Response Team Center, 1997b.

_____, *Risk Assessment Guidance for Superfund, Volume I: Human Health Evaluation Manual, Part D*, Washington, D.C.: EPA, 1998.

_____, *Risk Assessment Guidance for Superfund, Supplement to RAGS Part A: Community Involvement in Superfund Risk Assessments*, Washington, D.C.: EPA, 1999.

_____, *Handbook on the Management of Ordnance and Explosives at Closed, Transferred, and Transferring Ranges*, report no. EPA 505-B-01-001, Washington, D.C.: EPA, 2001.

Federal Aviation Administration (FAA), *Safety Risk Management*, order 8040.4, Washington, D.C.: Department of Transportation, Federal Aviation Administration, 1998.

_____, *System Safety Handbook: Practices and Guidelines for Conducting System Safety Engineering and Management*, Washington, D.C.: Department of Transportation, Federal Aviation Administration, 2000.

Fischhoff, Baruch, Sarah Lichtenstein, Paul Slovic, Stephen Derby, and Ralph L. Keeney, *Acceptable Risk*, New York: Cambridge University Press, 1981.

General Accounting Office (GAO), *Environmental Liabilities: DoD Training Range Cleanup Cost Estimates Are Likely Understated*, report no. GAO-01-479, Washington, D.C.: General Accounting Office, 2001.

Graham, John D., "Historical Perspective on Risk Assessment in the Federal Government," *Toxicology*, Vol. 102, Nos. 1–2, 1995, pp. 29–52.

Haimes, Yacov, *Risk Modeling Assessment, and Management*, New York: John Wiley & Sons, 1998.

Jaffe, Harry, "Spring Valley as Ground Zero," *Washingtonian*, Vol. 33, No. 3, 2003, pp. 121–129.

Jenkins, Thomas F., Alan D. Hewitt, Marianne E. Walsh, Thomas A. Ranney, Judy C. Pennington, Sonia Thiboutot, Guy Amplemen, and Martin H. Stutz, "Explosives Contamination at DoD Firing Ranges," paper delivered at The UXO/Countermine Forum, Orlando, FL, September 3–6, 2002.

Jenni, Karen E., Miley W. Merkhofer, and Carol Williams, "The Rise and Fall of a Risk-Based Priority System: Lessons from DOE's Environmental Restoration Priority System," *Risk Analysis*, Vol. 15, No. 3, 1995, pp. 397–410.

Kamins, Milton, *A Reliability Review of the Reactor Safety Study*, Santa Monica, CA: RAND, P-5413, 1975.

Keeney, Ralph L., and Howard R. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, New York: John Wiley & Sons, 1976.

Lauritzen, Erik Krogh, "The Challenge of Demilitarisation and Disposal of Ammunition," *Military Technology*, Vol. 25, No. 7, 2001, pp. 34–39.

Lewis, Harold W., et al., *Risk Assessment Review Group Report: A Critique of the Rasmussen Report* (WASH-1400), NUREG/CR-0400, Washington, D.C.: Nuclear Regulatory Commission, 1978.

MacDonald, Jacqueline A., "Cleaning Up Unexploded Ordnance," *Environmental Science & Technology*, Vol. 35, No. 17, 2001, pp. 372A–376A.

Malcolm Pirnie, *Review of Existing OE Risk Assessment Methods*, Baltimore, MD: Malcolm Pirnie, Inc., 2001.

Maly, Mary Ellen, "How Large Is the Scope? The Army's Range Inventory Program," paper presented at The UXO/Countermine Forum, Orlando, FL, September 3–6, 2002.

Melchers, R. E., "Probabilistic Systems Reliability," in D. Blockley (ed.), *Engineering Safety*, New York: McGraw Hill, 1992, pp. 117–137.

Mendez-Cruz, Carmen M., *Understanding the Issues Associated with Estimating the Cost of UXO Remediation on Closed, Transferred, and Transferring Military Training Ranges*, master's thesis, University of Wisconsin-Madison, December 2002.

Monterey Peninsula Herald, "Dud Kills Boy of 13," April 26, 1943, p. 1.

_____, "East Monterey Boy Loses Legs in Shell Explosion at Ord," April 18, 1949, p. 1.

_____, "Boy, Girl, 9, Wounded by Exploding Grenade," September 7, 1976, p. 1.

Morgan, M. Granger, "The Neglected Art of Bounding Analysis," *Environmental Science & Technology*, Vol. 35, No. 7, 2001, pp. 162A–164A.

Morgan, M. Granger, and Max Henrion, *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, New York: Cambridge University Press, 1990.

National Aeronautics and Space Administration (NASA), *Fault Tree Handbook with Aerospace Applications*, Washington, D.C.: NASA, Office of Safety and Mission Assurance, 2002a.

_____, *Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners*, Washington, D.C.: NASA Office of Safety and Mission Assurance, 2002b.

_____, *Risk Management Procedures and Guidelines*, Washington, D.C.: NASA Office of Safety and Mission Assurance, 2002c.

National Research Council, *Risk Assessment in the Federal Government*, Washington, D.C.: National Academy Press, 1983.

_____, *Alternatives for Ground Water Cleanup*, Washington, D.C.: National Academy Press, 1994a.

_____, *Ranking Hazardous Waste Sites for Remedial Action*, Washington, D.C.: National Academy Press, 1994b.

_____, *Understanding Risk: Informing Decisions in a Democratic Society*, Washington, D.C.: National Academy Press, 1996.

_____, *Innovations in Groundwater and Soil Cleanup*, Washington, D.C.: National Academy Press, 1997a.

_____, *Risk Assessment and Management at Deseret Chemical Depot and Tooele Chemical Agent Disposal Facility*, Washington, D.C.: National Academy Press, 1997b.

Nielson, Greg, and Sherri Anderson-Hudgins, "Spring Valley, Washington, D.C.: Formerly Utilized Defense Site," paper presented at The UXO/Countermine Forum, Orlando, FL, September 3–6, 2002.

North County Times, "Military Bomb Found in TierraSanta," April 1, 2001, <http://www.nctimes.net/news/2001/20010401/www.html>.

Nuclear Regulatory Commission (NRC), *Reactor Safety Study: An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants*, report WASH-1400, Washington, D.C.: NRC, 1975.

_____, *PRA Procedures Guide: A Guide to Performance of PRAs for Nuclear Power Plants*, Washington, D.C.: NRC, 1983.

_____, "Use of Probabilistic Risk Assessment Methods in Nuclear Activities: Final Policy Statement," *Federal Register*, Vol. 60, No. 158, August 16, 1995, pp. 42, 622–642, 629.

_____, *An Approach for Using Probabilistic Risk Assessment in Risk-Informed Decisions on Plant-Specific Changes to the Licensing Basis: Regulatory Guide*, Washington, D.C.: NRC, 1998a.

_____, *Use of Probabilistic Risk Assessment in Plant-Specific, Risk-Informed Decisionmaking: General Guidance*, Washington, D.C.: NRC, 1998b.

Nuclear Waste Technical Review Board, "Letter Report to the Congress and Secretary of Energy, Attachment 1, Part II," Washington, D.C.: U.S. Department of Energy, 2002.

Packer, Bonnie, and Kurt O. Thomsen, "Preliminary UXO Corrosion Data from Massachusetts Military Reservation," paper presented at The UXO/Countermine Forum, Orlando, FL, September 3–6, 2002.

Paustenbach, Dennis J., "The Practice of Health Risk Assessment in the United States (1975–1995): How the U.S. and Other Countries Can Benefit from that Experience," *Human and Ecological Risk Assessment*, Vol. 1, No. 1, 1995, pp. 29–79.

President's Commission on the Accident at Three Mile Island, *The Need for Change: The Legacy of TMI*, Washington, D.C.: President's Commission on the Accident at Three Mile Island, 1979.

Presidio of Monterey Directorate of Environmental and Natural Resources Management, *Prescribed Burns May Be Needed to Protect Public, Cleanup Workers, and Rare Habitat*, Community Bulletin #1, Presidio of Monterey, California, 2001.

Rasmussen, Norman C., "The Application of Probabilistic Risk Assessment Techniques to Energy Technologies," *Annual Review of Energy*, Vol. 6, 1981, pp. 123–138.

Rechard, Robert P., "Historical Relationship Between Performance Assessment for Radioactive Waste Disposal and Other Types of Risk Assessment," *Risk Analysis*, Vol. 19, No. 5, 1999, pp. 753–807.

Rogovin, Mitchell, and George T. Frampton, *Three Mile Island: A Report to the Commissioners and to the Public*, Washington, D.C.: Nuclear Regulatory Commission, 1980.

Shafer, Glenn, *The Art of Causal Conjecture*, Cambridge, MA: MIT Press, 1996.

Teachman, George, and Steve Getlein, "Range Response: Natural and Cultural Resources Issues," *Federal Facilities Environmental Journal*, Summer 2002, pp. 113–129.

References 151

Velleman, Paul F., and Leland Wilkinson, "Nominal, Ordinal, Interval, and Ratio Typologies Are Misleading," *The American Statistician*, Vol. 47, No. 1, 1993, pp. 65-72.